GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:49:36 ; Search time 10.5069 Seconds
(without alignments)
747.688 Million cell updates/sec

Title: US-09-654-652A-2
Perfect score: 1439
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....PNSSSVDKLAAALEHHHHHH 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1103	76.7	228	3	US-09-286-690-12	Sequence 12, Appl
2	245.5	17.1	242	3	US-09-286-690-8	Sequence 8, Appli
3	244	17.0	239	1	US-08-103-998-2	Sequence 2, Appli
4	231.5	16.1	279	3	US-09-286-690-9	Sequence 9, Appli
5	229.5	15.9	243	3	US-09-286-690-10	Sequence 10, Appl
6	225	15.6	238	3	US-09-286-690-7	Sequence 7, Appli
7	222.5	15.5	242	3	US-09-286-690-11	Sequence 11, Appl
8	220.5	15.3	237	1	US-08-103-998-4	Sequence 4, Appli
9	189	13.1	245	3	US-09-286-690-2	Sequence 2, Appli
10	154.5	10.7	306	2	US-08-824-707-2	Sequence 2, Appli
11	145	10.1	263	4	US-09-159-106-2	Sequence 2, Appli
12	143	9.9	435	4	US-09-159-106-11	Sequence 11, Appl
13	140.5	9.8	303	4	US-09-159-106-13	Sequence 13, Appl
14	133.5	9.3	276	2	US-08-712-072C-4	Sequence 4, Appli
15	118	8.2	321	2	US-08-712-072C-3	Sequence 3, Appli
16	117.5	8.2	289	2	US-08-580-545B-4	Sequence 4, Appli
17	117.5	8.2	289	4	US-09-262-653A-4	Sequence 4, Appli
18	117.5	8.2	289	4	US-08-867-484A-2	Sequence 2, Appli
19	117	8.1	620	4	US-09-269-731-6	Sequence 6, Appli
20	116.5	8.1	654	1	US-08-392-828C-2	Sequence 2, Appli
21	116.5	8.1	654	3	US-09-330-945-2	Sequence 2, Appli
22	115	8.0	262	1	US-08-392-828C-37	Sequence 37, Appl
23	115	8.0	262	3	US-09-330-945-37	Sequence 37, Appl
24	113	7.9	285	2	US-08-712-072C-5	Sequence 5, Appli
25	113	7.9	545	4	US-09-269-731-8	Sequence 8, Appli
26	110.5	7.7	295	1	US-08-295-657-3	Sequence 3, Appli
27	110	7.6	287	4	US-08-640-737-8	Sequence 8, Appli
28	109.5	7.6	289	4	US-08-640-737-2	Sequence 2, Appli
29	105	7.3	422	2	US-08-712-072C-2	Sequence 2, Appli
30	102.5	7.1	395	2	US-08-404-531B-3	Sequence 3, Appli
31	102.5	7.1	395	3	US-08-476-900A-3	Sequence 3, Appli
32	102.5	7.1	395	3	US-08-488-546A-3	Sequence 3, Appli
33	102.5	7.1	1580	4	US-08-726-320-1	Sequence 1, Appli
34	102.5	7.1	1580	4	US-09-208-716-1	Sequence 1, Appli
35	102.5	7.1	1581	4	US-08-726-320-3	Sequence 3, Appli
36	102.5	7.1	1581	4	US-09-208-716-3	Sequence 3, Appli
37	95.5	6.6	1581	2	US-08-404-531B-6	Sequence 6, Appli
38	95.5	6.6	1581	3	US-08-476-900A-6	Sequence 6, Appli
39	95.5	6.6	1581	3	US-08-488-546A-6	Sequence 6, Appli
40	95.5	6.6	1581	4	US-08-726-320-4	Sequence 4, Appli
41	95.5	6.6	1581	4	US-09-208-716-4	Sequence 4, Appli
42	95.5	6.6	1582	2	US-08-404-531B-9	Sequence 9, Appli
43	95.5	6.6	1582	3	US-08-476-900A-9	Sequence 9, Appli
44	95.5	6.6	1582	3	US-08-488-546A-9	Sequence 9, Appli
45	95.5	6.6	1582	4	US-08-726-320-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-286-690-12
; Sequence 12, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Fibrobacter succinogenes
US-09-286-690-12

Query Match 76.7%; Score 1103; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 3.3e-105;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84

Qy      62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144

Qy     122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204

Qy     182 FTLDWTDNFDTFDGSRWGKGDWTF 205
           ||||||||||||||||||||||||
Db     205 FTLDWTDNFDTFDGSRWGKGDWTF 228
```

RESULT 2
US-09-286-690-8
; Sequence 8, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:

```
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-286-690-8

  Query Match             17.1%;  Score 245.5;  DB 3;  Length 242;
  Best Local Similarity   35.8%;  Pred. No. 2.7e-17;
  Matches   59;  Conservative   24;  Mismatches   71;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |||:
Db      87 FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFLGKDT 143

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |   |
Db     144 TKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATNQ 199

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           :    |   :  |||:      |:| ::    ||:   :||:  |
Db     200 IPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 3
US-08-103-998-2
; Sequence 2, Application US/08103998
; Patent No. 5470725
;  GENERAL INFORMATION:
;    APPLICANT:  Borriss, Rainer
;    APPLICANT:  Hofemeister, Jurgen
;    APPLICANT:  Thomsen, Karl Kristian
;    APPLICANT:  Olsen, Ole
;    APPLICANT:  Von Wettstein, Dietrich
;    TITLE OF INVENTION:  A Thermostable (1,3-1,4)-beta-Glucanase
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  1800 Diagonal Road, P.O. Box 299
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/103,998
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/773,652
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30307/123
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-9300
;      TELEFAX:  (703) 683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  239 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-103-998-2

  Query Match             17.0%;  Score 244;  DB 1;  Length 239;
  Best Local Similarity   35.5%;  Pred. No. 3.8e-17;
  Matches   61;  Conservative   26;  Mismatches   69;  Indels   16;  Gaps    8;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |  |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      80 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 136

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |   |     ||  ::   | : |||| :| | |::| |||  :: |
Db     137 GKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT 192

Qy     123 EGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
               :|:  | | :  |||:      |:| :: :  ||:   :||  |||
Db     193 ---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 237


RESULT 4
US-09-286-690-9
; Sequence 9, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 279
;   TYPE: PRT
;   ORGANISM: Clostridium thermocellum
US-09-286-690-9

  Query Match             16.1%;  Score 231.5;  DB 3;  Length 279;
  Best Local Similarity   35.3%;  Pred. No. 9.1e-16;
  Matches   61;  Conservative   23;  Mismatches   70;  Indels   19;  Gaps    7;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           ::: |:  |   | :      || :| ||| |   | ||| | |   |   | || |:|
Db      82 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 138

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || |||:    | |      |   :|  | :   | | ||||| || |:|: : |||::
Db     139 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 195

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVY 169
           | :       |:  | | :  |||       |:|::|  : ||     :|| |
Db     196 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPLQAEYEYVKYY 243


RESULT 5
US-09-286-690-10
; Sequence 10, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
```

```
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Bacillus licheniformis
US-09-286-690-10

  Query Match             15.9%;  Score 229.5;  DB 3;  Length 243;
  Best Local Similarity   33.7%;  Pred. No. 1.2e-15;
  Matches   57;  Conservative   25;  Mismatches   76;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :|  || |   | ||| | |   :   || || |:||| |
Db      84 SYNKFDCGENRSVQTYGYGLYEVNMKPAKNVGIVSSFFTYTGPT---DGTPWDEIDIEFL 140

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |   |     ||   :   |  ::||| :| || ::| |||| :: |
Db     141 GKDTTKVQFNYYTNGVG---NHEKIVNLGFDAANSYHTYAFDWQPNSIKWYVDGQ-LKHT 196

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
             |:    |   :  |||:      |:| ::    || : ::||:  |
Db     197 ATTQIPQTPGK--IMMNLWNGAGVDEWLGSYN-GVTPLSRSLHWVRYTK 242


RESULT 6
US-09-286-690-7
; Sequence 7, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Bacillus polymyxa
US-09-286-690-7

  Query Match             15.6%;  Score 225;  DB 3;  Length 238;
  Best Local Similarity   34.5%;  Pred. No. 3.3e-15;
  Matches   58;  Conservative   23;  Mismatches   71;  Indels   16;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :     || :|  || |  :| ||| | |   |    |  | |:||| |||:
Db      83 FDCGEYRSTNNYGYGLYEVSMKPAKNTGIVSSFFTYTGPSH---GTQWDEIDIEFLGKDT 139

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |     ||   :   |  :|||| :| | |::| |||  :: |
Db     140 TKVQFNYYTNGVGGH---EKIINLGFDASTSFHTYAFDWQPGYIKWYVDG-VLKHT---A 192

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
             :|:  | | :  |||:     :|:| :: :  ||:   :||   |||
Db     193 TTNIPSTPGKIMMNLWNGTGVDSWLGSYNGAN-PLYAEYDWV---KYT 236


RESULT 7
US-09-286-690-11
; Sequence 11, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Clostridium thermocellum
US-09-286-690-11

  Query Match             15.5%;  Score 222.5;  DB 3;  Length 242;
  Best Local Similarity   35.6%;  Pred. No. 6.2e-15;
  Matches   58;  Conservative   22;  Mismatches   64;  Indels   19;  Gaps    7;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
             ::: |:  |   | :      || :| ||| |   | ||| | |   |   |  || |:|
Db      80 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 136

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || |||:    | |      |    :|  | :   | | ||||| || |:|: : |||::
Db     137 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 193

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPL 159
           | :       |:  | | :  |||       |:|::|  : ||
Db     194 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPL 231


RESULT 8
US-08-103-998-4
; Sequence 4, Application US/08103998
; Patent No. 5470725
;  GENERAL INFORMATION:
;    APPLICANT:  Borriss, Rainer
;    APPLICANT:  Hofemeister, Jurgen
;    APPLICANT:  Thomsen, Karl Kristian
;    APPLICANT:  Olsen, Ole
;    APPLICANT:  Von Wettstein, Dietrich
;    TITLE OF INVENTION:  A Thermostable (1,3-1,4)-beta-Glucanase
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  1800 Diagonal Road, P.O. Box 299
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/103,998
;      FILING DATE:
;      CLASSIFICATION:  435
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/773,652
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30307/123
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-9300
;      TELEFAX:  (703) 683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  237 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-103-998-4

  Query Match             15.3%;  Score 220.5;  DB 1;  Length 237;
  Best Local Similarity   33.9%;  Pred. No. 9.6e-15;
  Matches   56;  Conservative   23;  Mismatches   75;  Indels   11;  Gaps    6;

Qy        7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
              | ||  :       || :|  || |  :| ||| |         | |   | |:||| |||:
Db       82 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138

Qy       67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
                 | |  |  ||       ||    :    |  |:|||  :| || ::| |||| :: |    |
Db      139 TKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATTQ 194

Qy      127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
              :     |    :   |||:        |:| ::     |::    :|:: |
Db      195 IPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYDWMRYRK 236


RESULT 9
US-09-286-690-2
; Sequence 2, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 245
;   TYPE: PRT
;   ORGANISM: Orpinomyces sp. PC-2
US-09-286-690-2

  Query Match             13.1%;  Score 189;  DB 3;  Length 245;
  Best Local Similarity   32.9%;  Pred. No. 1.7e-11;
  Matches   56;  Conservative   16;  Mismatches   76;  Indels   22;  Gaps    8;

Qy        5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
              :| ||     ||  | :        || |:  ||       | ||| | |     |     ||   | |:|
Db       83 RDGSG---YTCGEYRTKNYYGYGMFQVNMKPIKNPGVVSSFFTYTGPS---DGTKWDEID 136

Qy       59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
              || || :        | |   |    |           |   | :     | | |||||  |   | : | |||
Db      137 IEFLGYDTTKVQFNYYTNGQGHH---EHIHYLGFDASQGFHTYGFFWARNSITWYVDGTA 193

Qy      119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWV 166
               |                 |:   |  |  :    |  |:       |:   |:  :   :  :|||
Db      194 VYTA----YDNIPDTPGKIMMNAWNGIGVDDWLRPFN-GRTNISAYYDWV 238


RESULT 10
US-08-824-707-2
; Sequence 2, Application US/08824707
; Patent No. 5919688
;  GENERAL INFORMATION:
;    APPLICANT:  Ferrer, Pau
;    APPLICANT:  Diers, Ivan
;    APPLICANT:  Hedegaard, Lisbeth
;    APPLICANT:  Halkier, Torben
;    APPLICANT:  Asenjo, Juan
;    APPLICANT:  Savva, Demitris
;    TITLE OF INVENTION:  No. 5919688el enzyme with beta-1,3-glucanase activity
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5919688o No. 5919688disk of No. 5919688th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/824,707
;      FILING DATE:  14-April-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gregg, Valeta A.
;      REGISTRATION NUMBER:  35,127
;      REFERENCE/DOCKET NUMBER:  4290.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  306 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-824-707-2

  Query Match             10.7%;  Score 154.5;  DB 2;  Length 306;
  Best Local Similarity   28.5%;  Pred. No. 8.1e-08;
  Matches   53;  Conservative   28;  Mismatches   68;  Indels   37;  Gaps   11;

Qy        7 FSGAELYTLEEV--QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI-E 60
              :: |  |  |      |  |:|: |||:::       |   |: ::      |: :  |  ||     |:|| |
Db      131 YTSARLTTQGNVQPQFGRIEARIQIPRGQGIWSAFWMV--GANLPD-TPWPTSGEIDIME 187

Qy       61 VLGKNP----GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 116
               :|  |      |:         :|  |     |  :          | | |   |||:|::|||   :  |  |||
Db      188 NVGNAPHEVHGTVHGPGYSGDNGIMGTYQHPQGWSFADD--FHTFGIDWTPGEITWLVDG 245

Qy      117 QEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQG 176
              ||  :        |                          :  ||   ||:    |  |   :| |  :      ||
Db      246 QEYHRVTTADVG----------------ANQWV--FDQ---PFFLILN-VAIGGQWPGNP 283

Qy      177 EGGSDF 182
              :   :  |
Db      284 DATTPF 289


RESULT 11
US-09-159-106-2
; Sequence 2, Application US/09159106
; Patent No. 6284509
```

```
; GENERAL INFORMATION:
;  APPLICANT: Ferrer, Pau
;  APPLICANT: Diers, Ivan
;  APPLICANT: Halkier, Torben
;  APPLICANT: Hedegaard, Lisbeth
;  TITLE OF INVENTION: An Enzyme With  -1,3-Glucanase
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 4693.204-US
;  CURRENT APPLICATION NUMBER: US/09/159,106
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: 0427/96
;  EARLIER FILING DATE: 1996-12-04
;  EARLIER APPLICATION NUMBER: 0885/96
;  EARLIER FILING DATE: 1996-08-23
;  EARLIER APPLICATION NUMBER: PCT/DK97/00160
;  EARLIER FILING DATE: 1997-04-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 263
;   TYPE: PRT
;   ORGANISM: Oerskovia xanthineolytica
US-09-159-106-2

  Query Match             10.1%;  Score 145;  DB 4;  Length 263;
  Best Local Similarity   25.9%;  Pred. No. 6.1e-07;
  Matches   49;  Conservative   29;  Mismatches   73;  Indels   38;  Gaps   10;

Qy      19 QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITG-- 76
           |||: |||:::    | :   |    ||    | ||   :|::: :| | |: : : |
Db      82 QYGRIEARIQIPRGQG-IWPAFWMLGGS--FPGTPWPSGEIDIM-ENVG-FEPHRVHGTV 136

Qy      77 -----KAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLT 131
                 |:  |    |    :    |||: ::| |  : | ||||:  :     |
Db     137 HGPGYSGGSGITGMYQHPQGWSFADTFHTFAVDWKPGEITWFVDGQQFHRVTRASVG--- 193

Qy     132 GTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS----DFTLDWT 187
                       : |||  ||:   | |  :| | |    ||  :| :       :|:
Db     194 -------------ANAWV--FDQ---PFFLILN-VAVGGQWPGYPDGTTQLPQQMKVDYV 234

Qy     188 DNFDTFDGS 196
           :|   ||
Db     235 RVYDNGSGS 243


RESULT 12
US-09-159-106-11
; Sequence 11, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
;  APPLICANT: Ferrer, Pau
;  APPLICANT: Diers, Ivan
;  APPLICANT: Halkier, Torben
;  APPLICANT: Hedegaard, Lisbeth
;  TITLE OF INVENTION: An Enzyme With  -1,3-Glucanase
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 4693.204-US
;  CURRENT APPLICATION NUMBER: US/09/159,106
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: 0427/96
;  EARLIER FILING DATE: 1996-12-04
;  EARLIER APPLICATION NUMBER: 0885/96
;  EARLIER FILING DATE: 1996-08-23
;  EARLIER APPLICATION NUMBER: PCT/DK97/00160
;  EARLIER FILING DATE: 1997-04-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
;   LENGTH: 435
;   TYPE: PRT
;   ORGANISM: Oerskovia xanthineolytica
US-09-159-106-11

  Query Match             9.9%;  Score 143;  DB 4;  Length 435;
  Best Local Similarity   24.6%;  Pred. No. 2.1e-06;
  Matches   62;  Conservative   29;  Mismatches   95;  Indels   66;  Gaps   14;

Qy      19 QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI-EVLGKNP----GSFQ 70
           |||: |||:::    | :   |    ||    | ||    |:|| | :|  |    |:
Db     134 QYGRIEARIQIPRGQG-IWPAFWMLGGS--FPGTPWPSSGEIDIMENVGFEPHRVHGTVH 190

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
             :| :|   |    |    :   |||: ::| |  : | ||||:  :     |
Db     191 GPGYSGGSGI--TGMYQHPQGWSFADTFHTFAVDWKPGEITWFVDGQQFHRVTRASVG-- 246

Qy     131 TGTQGLRFNLWSSE----------SAAWVGQFD-ESKLPLFQFINWVKVYKYTPGQGEG 178
                 | |  :              | |  |  ::||    :::|:||    |
Db     247 -------ANAWVFDQPFFLILNVAVGGQWPGYPDGTTQLPQQMKVDYVRVYDNGSGSSSP 299

Qy     179 GSDFT-----------------LDWTDNFD-------TFDGS---RWGKGDWTFDGNRV 210
           |:  |                 : | |  |       |  |:    | :|    ||  |
Db     300 GNPGTGLPTGTGAVRAANGMCVDVPWADPTDGNPVQIVTCSGNAAQTWTRGS---DGT-V 355

Qy     211 DLTDKNIYSRDG 222
             | :  |||
Db     356 RALGKCLDVRDG 367


RESULT 13
US-09-159-106-13
; Sequence 13, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
;  APPLICANT: Ferrer, Pau
;  APPLICANT: Diers, Ivan
;  APPLICANT: Halkier, Torben
;  APPLICANT: Hedegaard, Lisbeth
;  TITLE OF INVENTION: An Enzyme With  -1,3-Glucanase
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 4693.204-US
;  CURRENT APPLICATION NUMBER: US/09/159,106
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: 0427/96
;  EARLIER FILING DATE: 1996-12-04
;  EARLIER APPLICATION NUMBER: 0885/96
;  EARLIER FILING DATE: 1996-08-23
;  EARLIER APPLICATION NUMBER: PCT/DK97/00160
;  EARLIER FILING DATE: 1997-04-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 303
;   TYPE: PRT
;   ORGANISM: Oerskovia xanthineolytica
US-09-159-106-13

  Query Match             9.8%;  Score 140.5;  DB 4;  Length 303;
  Best Local Similarity   27.4%;  Pred. No. 2.2e-06;
  Matches   52;  Conservative   25;  Mismatches   74;  Indels   39;  Gaps   11;

Qy      19 QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI-EVLGKNP----GSFQ 70
           |||: |||:::    | :   |    ||    | ||    |:|| | :|  |    |:
Db     134 QYGRIEARIQIPRGQG-IWPAFWMLGGS--FPGTPWPSSGEIDIMENVGFEPHRVHGTVH 190

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
             :| :|   |    |    :   |||: ::| |  : | ||||:  :     |
Db     191 GPGYSGGSGI--TGMYQHPQGWSFADTFHTFAVDWKPGEITWFVDGQQFHRVTRASVG-- 246

Qy     131 TGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS----DFTLDW 186
                       : |||  ||:   | |  :| | |    ||  :| :       :|:
Db     247 -------------ANAWV--FDQ---PFFLILN-VAVGGQWPGYPDGTTQLPQQMKVDY 286

Qy     187 TDNFDTFDGS 196
           :|   ||
Db     287 VRVYDNGSGS 296
```

```
RESULT 14
US-08-712-072C-4
; Sequence 4, Application US/08712072C
; Patent No. 5925541
;  GENERAL INFORMATION:
;    APPLICANT:  Jack Goldstein, Alex Zhu and Lin Leng
;    TITLE OF INVENTION:  ENDO-BETA-GALACTOSIDASE
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amster, Rothstein & Ebenstein
;      STREET:  90 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH 1.44 Mb STORAGE DISKETTE
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/712,072C
;      FILING DATE:  11-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Elizabeth A.
;      REGISTRATION NUMBER:  39,911
;      REFERENCE/DOCKET NUMBER:  63475/97
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-5995
;      TELEFAX:  (212) 286-0854 or 286-0082
;      TELEX:  TWX 710-581-4766
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  gub, Rhodothermus marinus
US-08-712-072C-4

  Query Match             9.3%;  Score 133.5;  DB 2;  Length 276;
  Best Local Similarity   23.7%;  Pred. No. 9.9e-06;
  Matches   44;  Conservative   34;  Mismatches   85;  Indels   23;  Gaps    8;

Qy       5 KDFSGAELYTLEEVQ--YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI 59
           :::: | | |  :    ||:|| | :: :  ||  ::::  :          |    |:||
Db      93 REYTSARLVTRGKASWTYGRFEIRARLPSGRGTWPAIWMLPDRQTYGSAY-WPDNGEIDI 151

Qy      60 -EVLGKNP----GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTV 114
            | :| ||    |:  :       | |:         | |   || | :|||| :|| |
Db     152 MEHVGFNPDVVHGTVHTKAYNHLLGTQRGGSIR---VPTARTDFHVYAIEWTPEEIRWFV 208

Qy     115 DGQEVRKTEGGQVSNLTG-------TQGLRFNLWSSESAAWVGQ--FDESKLPLFQFINW 165
           |     :    ::::          |     :  :   || ||   |    |    :::
Db     209 DDSLYYRFPNERLTDPEADWRHWPFDQPFHLIMNIAVGGAWGGQQGVDPEAFPAQLVVDY 268

Qy     166 VKVYKY 171
           |:||::
Db     269 VRVYRW 274


RESULT 15
US-08-712-072C-3
```

```
; Sequence 3, Application US/08712072C
; Patent No. 5925541
;  GENERAL INFORMATION:
;    APPLICANT:  Jack Goldstein, Alex Zhu and Lin Leng
;    TITLE OF INVENTION:  ENDO-BETA-GALACTOSIDASE
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amster, Rothstein & Ebenstein
;      STREET:  90 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH 1.44 Mb STORAGE DISKETTE
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/712,072C
;      FILING DATE:  11-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Elizabeth A.
;      REGISTRATION NUMBER:  39,911
;      REFERENCE/DOCKET NUMBER:  63475/97
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-5995
;      TELEFAX:  (212) 286-0854 or 286-0082
;      TELEX:  TWX 710-581-4766
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  321 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  el3b, Bacillus circulars
US-08-712-072C-3

  Query Match             8.2%;  Score 118;  DB 2;  Length 321;
  Best Local Similarity   20.5%;  Pred. No. 0.00048;
  Matches   40;  Conservative   31;  Mismatches   80;  Indels   44;  Gaps    6;

Qy       7 FSGAELYTLEE--VQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI-EVLG 63
           :|  :: | ::  ::||: : | |:    |   ::::    |          |:|: |  |
Db     140 YSSGKINTKDKLSLKYGRVDFRAKLPTGDGVWPALWMLPKDSVYGTWAASGEIDVMEARG 199

Qy      64 KNPGSFQSNIITG---KAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120
           : |||    |  |           : |          :| | : |  : ::| |||:
Db     200 RLPGSVSGTIHFGGQWPVNQSSGGDYHFPEGQTFANDYHVYSVVWEEDNIKWYVDGKFFY 259

Qy     121 KTEGGQVSNLTGTQGLRFNLWSSESA-----------------AWVGQFD------ESK 156
           |    |               | | :|                 |  | ||       |
Db     260 KVTNQQ-------------WYSTAAPNNPNAPFDEPFYLIMNLAVGGNFDGGRTPNASD 305

Qy     157 LPLFQFINWVKVYKY 171
           :|    :::|:|||:
Db     306 IPATMQVDYVRVYKF 320


Search completed: January  9, 2003, 12:06:33
Job time : 11.5069 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:14:27 ; Search time 12.6701 Seconds
(without alignments)
2025.857 Million cell updates/sec

Title: US-09-654-652A-2
Perfect score: 1439
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....PNSSSVDKLAAALEHHHHHH 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1340	93.1	349	2	A44507	licheninase (EC 3.
2	266.5	18.5	851	2	H84053	endo-beta-1,3-1,4
3	265.5	18.5	252	2	A48378	licheninase (EC 3.
4	247.5	17.2	276	2	I40453	licheninase (EC 3.
5	245.5	17.1	242	1	LXBS	licheninase (EC 3.
6	242.5	16.9	802	2	A36910	xylanase, beta(1,3
7	238.5	16.6	239	1	A29091	licheninase (EC 3.
8	237	16.5	334	1	S23498	licheninase (EC 3.
9	229.5	15.9	243	1	S15388	licheninase (EC 3.
10	226	15.7	237	1	S11927	licheninase (EC 3.
11	225	15.6	238	1	S19012	licheninase (EC 3.
12	222.5	15.5	242	2	JS0611	endo-1,3(4)-beta-g
13	221	15.4	246	2	D97245	endo-1,3(4)-beta-g
14	206	14.3	263	2	AB3055	endo-1,3-1,4-beta-
15	206	14.3	289	2	C98231	endo-1,3-1,4-beta-
16	197	13.7	269	2	H95976	endo-beta-1,3-1,4-
17	190.5	13.2	302	2	G84053	hybrid-endo-beta-1
18	182.5	12.7	507	2	S64507	probable membrane
19	180	12.5	268	2	S34804	exoK protein - Rhi
20	172	12.0	282	2	T02354	xyloglucan endo-1,
21	171.5	11.9	277	2	B85354	hypothetical prote
22	169.5	11.8	282	2	A85354	hypothetical prote
23	169	11.7	642	2	B72428	laminarinase - The
24	163	11.3	286	2	S71225	xyloglucan endo-1,
25	159	11.0	287	2	T04236	xyloglucan endo-1,
26	154.5	10.7	277	2	S71222	xyloglucan endo-1,
27	150.5	10.5	269	2	S61555	xyloglucan endo-1,
28	147	10.2	467	2	S30839	UTR2 protein - yea
29	146.5	10.2	284	2	T52097	xyloglucan endo-1,
30	142	9.9	422	2	S48564	probable membrane
31	141.5	9.8	289	2	T06166	xyloglucan endotra
32	140.5	9.8	280	2	T02090	xyloglucan endo-1,
33	140.5	9.8	310	2	A86239	protein T10024.17
34	140	9.7	305	2	G84568	probable xylogluca
35	139.5	9.7	286	2	T06202	xyloglucan endo-1,
36	133.5	9.3	286	2	S48201	licheninase (EC 3.
37	132	9.2	295	2	T10210	xyloglucan endo-1,
38	131.5	9.1	292	2	T06201	xyloglucan endo-1,
39	129.5	9.0	1324	2	T18265	endo-1,3(4)-beta-g
40	128	8.9	283	2	T07678	xyloglucan endo-1,
41	127.5	8.9	269	2	T05895	xyloglucan endo-1,
42	127	8.8	292	2	T04514	xyloglucan endo-1,
43	125.5	8.7	277	2	JE0156	end-xyloglucan tra
44	125	8.7	301	2	C87296	beta-glucanase [im
45	124	8.6	299	2	D84519	probable endoxylog

ALIGNMENTS

RESULT 1
A44507
licheninase (EC 3.2.1.73) - Fibrobacter succinogenes
C;Species: Fibrobacter succinogenes
C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 08-Oct-1999
C;Accession: A44507
R;Teather, R.M.; Erfle, J.D.
J. Bacteriol. 172, 3837-3841, 1990
A;Title: DNA sequence of a Fibrobacter succinogenes mixed-linkage beta-glucanase (1,3
A;Reference number: A44507; MUID:90299807; PMID:2193918
A;Accession: A44507
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TEA>
A;Cross-references: EMBL:M33676; NID:g148575; PIDN:AAA24896.1; PID:g148576
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 93.1%; Score 1340; DB 2; Length 349;
Best Local Similarity 99.6%; Pred. No. 1.5e-104;
Matches 250; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61
           ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84

Qy      62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144

Qy     122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204

Qy     182 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 264

Qy     242 RDDEPAPNSSS 252
           ||||||| |||
Db     265 RDDEPAPQSSS 275
```

RESULT 2
H84053
endo-beta-1,3-1,4 glucanase (licheninase) bglS [imported] - Bacillus halodurans (stra
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: H84053
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: H84053
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-851 <STO>
A;Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB06951.1; GSPDB:GN0
A;Experimental source: strain C-125
C;Genetics:
A;Gene: bglS

Query Match 18.5%; Score 266.5; DB 2; Length 851;
Best Local Similarity 29.7%; Pred. No. 3e-14;
Matches 82; Conservative 31; Mismatches 108; Indels 55; Gaps 12;

```
Qy       4 AKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLG 63
           |  :   || | :  ||| ||  || |   ||||| | |  |    || || |:||| ||
Db      99 APPYKAGELRTNDFYQYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFLG 157

Qy      64 KNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
           |:    | |  |    |    :| :: :   | ::|:||  ||  : : | |:|: |
Db     158 KDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA- 213

Qy     124 GGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP----GQGE 177
             |:  | | :  |||       | | ||    ||     ||   :|||     : |
Db     214 ---TENIPQTPQKIMMNLWPGVGVDEWTGVFDGDNTPLHADYEWV---RYTPLEALDEEE 267

Qy     178 GGSDFTLDWT----------------------DNFDTFDGSRWG-KGDWTFDGNRVDLT 213
             | |   :                       : |:||:   |     || :|   : |
Db     268 SGDDNEEEPVEEVEEEPAEDEEVSVRPIGSAIYETFNTFNEDIWSIAHGWT-NGQMFNAT 326

Qy     214 --DKNIYSRDGMLILALTRKGQESFNGQVPRDDEPA 247
             |  |   :|::  |: ::           ||| |
Db     327 WYDSQITFSNGIMRFAIDKE----------DDEEA 351
```

RESULT 3
A48378
licheninase (EC 3.2.1.73) - Bacillus brevis
N;Alternate names: lichenase
C;Species: Bacillus brevis
C;Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 07-May-1999
C;Accession: A48378
R;Louw, M.E.; Reid, S.J.; Watson, T.G.
Appl. Microbiol. Biotechnol. 38, 507-513, 1993
A;Title: Characterization, cloning and sequencing of a thermostable endo-(1,3-1,4) beta-
A;Reference number: A48378; MUID:93159752; PMID:7763386
A;Accession: A48378
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-252 <LOU>
A;Note: sequence extracted from NCBI backbone (NCBIN:124672, NCBIP:124673)
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 18.5%; Score 265.5; DB 2; Length 252;
Best Local Similarity 35.8%; Pred. No. 7.9e-15;
Matches 62; Conservative 24; Mismatches 74; Indels 13; Gaps 6;

```
Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           ||:::   || | :   || ||  || |   ||||| | |  |    || || |:||| |
Db      90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |    |    :| :: :   | ::|:||  ||  : : | |:|: |
Db     149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy     123 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             |:  | | :  |||       | | ||    |:: : :||   :|||
Db     206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTP 251
```

RESULT 4
I40453
licheninase (EC 3.2.1.73) - Bacillus sp.

C;Species: Bacillus sp.
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 22-Jun-1999
C;Accession: I40453; S32688
R;Tabernero, C.; Coll, P.M.; Fernandez-Abalos, J.M.; Perez, P.; Santamaria, R.I.
Appl. Environ. Microbiol. 60, 1213-1220, 1994
A;Title: Cloning and DNA sequencing of bgaA, a gene encoding an endo-beta-1,3-1,4-glu
A;Reference number: I40453; MUID:94288605; PMID:7517127
A;Accession: I40453
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-276 <RES>
A;Cross-references: EMBL:Z12151; NID:g296931; PIDN:CAA78135.1; PID:g296932
C;Genetics:
A;Gene: bgaI
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase

Query Match 17.2%; Score 247.5; DB 2; Length 276;
Best Local Similarity 36.1%; Pred. No. 2.8e-13;
Matches 61; Conservative 20; Mismatches 75; Indels 13; Gaps 6;

```
Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           :   || | :   || ||  || | ::||||| | |  |    :  || |:||| |||:
Db      96 YKAGELRTNDFYHYGLFEVSMKPAKSTGTVSSFFTY-TGPWDWENDPWDEIDIEFLGKDT 154

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |    :| :| :   |   |:||  || |  :|| |:|: |
Db     155 TKIQFNYFTNGVGG---NEHYHELGFDAADDFNTYAFEWRPESIRWFVNGELVHTA---- 207

Qy     127 VSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             |:  | | :  |||       | |:|:    |:    :||   ||||
Db     208 TENIPQTPQKIMMNLWPGIGVDGWTGRFNGEDTPVVTQYDWV---KYTP 253
```

RESULT 5
LXBS
licheninase (EC 3.2.1.73) precursor - Bacillus subtilis
N;Alternate names: 1,3-1,4-beta-D-glucan 4-gluconohydrolase; beta-glucanase; lichenas
N;Contains: licheninase E-1; licheninase E-2
C;Species: Bacillus subtilis
C;Date: 30-Jun-1988 #sequence_revision 13-Mar-1998 #text_change 16-Jun-2000
C;Accession: I40370; B69594; A22914; A90026; A90027; JU0110
R;Wolf, M.; Geczi, A.; Simon, O.; Borriss, R.
Microbiology 141, 281-290, 1995
A;Title: Genes encoding xylan and beta-glucan hydrolysing enzymes in Bacillus subtili
A;Reference number: I40370; MUID:95219081; PMID:7704256
A;Accession: I40370
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-242 <RES>
A;Cross-references: EMBL:Z46862; NID:g599673; PIDN:CAA86922.1; PID:g599674
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: B69594
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-242 <KUN>
A;Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB15943.1; PID:g26364
A;Experimental source: strain 168

```
R;Murphy, N.; McConnell, D.J.; Cantwell, B.A.
Nucleic Acids Res. 12, 5355-5367, 1984
A;Title: The DNA sequence of the gene and genetic control sites for the excreted B. subt
A;Reference number: A93526; MUID:84272222; PMID:6087283
A;Accession: A22914
A;Molecule type: DNA
A;Residues: 1-203,'L',205-242 <MUR>
A;Cross-references: EMBL:X00754; NID:g39818; PIDN:CAA25328.1; PID:g685236
A;Experimental source: strain C120
A;Note: the authors translated the codon CAA for residue 29 as Lys and CCA for residue 8
R;Tezuka, H.; Yuuki, T.; Yabuuchi, S.
Agric. Biol. Chem. 53, 2335-2339, 1989
A;Title: Construction of a beta-glucanase hyperproducing Bacillus subtilis using the clo
A;Reference number: A90026
A;Accession: A90026
A;Molecule type: DNA
A;Residues: 1-23,'S',25-82,'S',84-242 <TEZ>
A;Cross-references: DDBJ:D00518; NID:g216243; PIDN:BAA00405.1; PID:g216244
A;Experimental source: strain Y-25, clone pLE100
R;Yuuki, T.; Tezuka, H.; Yabuuchi, S.
Agric. Biol. Chem. 53, 2341-2346, 1989
A;Title: Purification and some properties of two enzymes from a beta-glucanase hyperprod
A;Reference number: A90027
A;Contents: annotation
A;Note: source was hyperproducing strain HL-25 with gene from strain Y-25
A;Note: the amino ends of the mature forms of E-1 and E-2 are pyroglutamic acid and glut
C;Genetics:
A;Gene: bglS
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin and c
C;Superfamily: licheninase
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation; p
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-242/Product: licheninase #status predicted <MAT>
F;29/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) (partial) #status

  Query Match             17.1%;  Score 245.5;  DB 1;  Length 242;
  Best Local Similarity   35.8%;  Pred. No. 3.5e-13;
  Matches   59;  Conservative   24;  Mismatches   71;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |||:
Db      87 FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFLGKDT 143

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |   |
Db     144 TKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATNQ 199

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           :   |   :  |||:      |:| ::    ||:   :||:  |
Db     200 IPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 6
A36910
xylanase, beta(1,3-1,4)-glucanase - Ruminococcus flavefaciens
C;Species: Ruminococcus flavefaciens
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
C;Accession: A36910
R;Flint, H.J.; Martin, J.; McPherson, C.A.; Daniel, A.S.; Zhang, J.X.
J. Bacteriol. 175, 2943-2951, 1993
A;Title: A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-glucanase domai
A;Reference number: A36910; MUID:93259938; PMID:8491715
A;Accession: A36910
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-802 <FLI>
A;Cross-references: GB:S61204; NID:g385910; PIDN:AAB26620.1; PID:g385911
A;Note: sequence extracted from NCBI backbone (NCBIN:131871, NCBIP:131872)
C;Superfamily: endo-1,4-beta-xylanase homology; Thermotoga xylanase A amino-terminal rep
F;42-239/Domain: endo-1,4-beta-xylanase homology <XYL>
F;259-401/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

  Query Match             16.9%;  Score 242.5;  DB 2;  Length 802;
  Best Local Similarity   34.9%;  Pred. No. 2.9e-12;
  Matches   58;  Conservative   21;  Mismatches   74;  Indels   13;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           :|| |  |     || :|  |:     | ||| | |   |   |  || |:|||:||||
Db     634 YSGGEFRTNNFYHYGYYECSMQAMKNDGVVSSFFTYTGPS---DDNPWDEIDIEILGKNT 690

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |     || : :   : :|:|||| :| |||: | |||:|| :
Db     691 TQVQFNYYTNGQGKH---EKLYDLGFDSSEAYHTYGFDWQPNYIAWYVDGREVYRA---- 743

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
             ::  | | :  | |   :   |:  |:  : ||     ||   |
Db     744 TQDIPKTPGKIMMNAWPGLTVDDWLKAFN-GRTPLTAHYQWVTYNK 788


RESULT 7
A29091
licheninase (EC 3.2.1.73) beta - Bacillus amyloliquefaciens
N;Alternate names: 1,3-1,4-beta-D-glucan 4-glucanohydrolase; beta-glucanase; lichenas
C;Species: Bacillus amyloliquefaciens
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29091
R;Hofemeister, J.; Kurtz, A.; Borriss, R.; Knowles, J.
Gene 49, 177-187, 1986
A;Title: The beta-glucanase gene from Bacillus amyloliquefaciens shows extensive homo
A;Reference number: A91564; MUID:87192007; PMID:3106158
A;Accession: A29091
A;Molecule type: DNA
A;Residues: 1-239 <HOF>
A;Cross-references: GB:M15674; NID:g143009; PIDN:AAA87323.1; PID:g143010
A;Experimental source: strain BE20/78
C;Genetics:
A;Gene: bglA
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match             16.6%;  Score 238.5;  DB 1;  Length 239;
  Best Local Similarity   34.3%;  Pred. No. 1.3e-12;
  Matches   58;  Conservative   26;  Mismatches   74;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   | :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      80 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 136

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:   | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
Db     137 GKDTTKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 192

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
             |:   |   :  |||:      |:| ::    |::   :|::  |
Db     193 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYDWMRYRK 238


RESULT 8
S23498
licheninase (EC 3.2.1.73) licB precursor - Clostridium thermocellum
N;Alternate names: beta-1,3-1,4-glucanase licB; lichenase licB
C;Species: Clostridium thermocellum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S23498; S22137
R;Schimming, S.; Schwarz, W.H.; Staudenbauer, W.L.
Eur. J. Biochem. 204, 13-19, 1992
A;Title: Structure of the Clostridium thermocellum gene licB and the encoded beta-1,3
1 cellulases.
A;Reference number: S23498; MUID:92155194; PMID:1740123
A;Accession: S23498
A;Molecule type: DNA
A;Residues: 1-334 <SCH>
A;Cross-references: EMBL:X63355; NID:g40697; PIDN:CAA44959.1; PID:g40698
C;Genetics:
A;Gene: licB
```

```
C;Superfamily: licheninase licB; Clostridium cellulase repeat homology
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-334/Product: licheninase #status predicted <MAT>
F;273-296/Domain: Clostridium cellulase repeat homology <CCR1>
F;308-331/Domain: Clostridium cellulase repeat homology <CCR2>

  Query Match             16.5%;  Score 237;  DB 1;  Length 334;
  Best Local Similarity   30.9%;  Pred. No. 2.7e-12;
  Matches   80;  Conservative   27;  Mismatches   94;  Indels   58;  Gaps   13;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           ::: |:  |   | :      || :| ||| |   | ||| | |   |   |  || |:|
Db      82 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 138

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || |||:    | |      |    :|  | :   | | ||||| || |:|: : |||::
Db     139 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 195

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVY------- 169
           | :       |:  | | :  |||       |:|::|  : ||     :|| |
Db     196 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPLQAEYEYVKYYPNGVPQD 250

Qy     170 ---------KYTP-----------GQGE-GGSDFT------LDWTDNFDTFDGSRWGKG 201
                   ||            | |    ||::      |   | |   | |
Db     251 NPTPTPTIAPSTPTNPNLPLKGDVNGDGHVNSSDYSLFKRYLLRVIDRFPVGDQS---VA 307

Qy     202 DWTFDGNRVDLTDKNIYSR 220
           |   || |:| ||  :  |
Db     308 DVNRDG-RIDSTDLTMLKR 325


RESULT 9
S15388
licheninase (EC 3.2.1.73) - Bacillus licheniformis
N;Alternate names: endo-beta-1,3-1,4-D-glucanase; lichenase
C;Species: Bacillus licheniformis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S15388
R;Lloberas, J.; Perez-Pons, J.A.; Querol, E.
Eur. J. Biochem. 197, 337-343, 1991
A;Title: Molecular cloning, expression and nucleotide sequence of the endo-beta-1,3-1,4-
A;Reference number: S15388; MUID:91224124; PMID:2026156
A;Accession: S15388
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-243 <LLO>
A;Cross-references: EMBL:X57279; NID:g39558; PIDN:CAA40547.1; PID:g39559
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match             15.9%;  Score 229.5;  DB 1;  Length 243;
  Best Local Similarity   33.7%;  Pred. No. 7.8e-12;
  Matches   57;  Conservative   25;  Mismatches   76;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :|  || |   | ||| | |   :   || || |:||| |
Db      84 SYNKFDCGENRSVQTYGYGLYEVNMKPAKNVGIVSSFFTYTGPT---DGTPWDEIDIEFL 140

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |   |     ||   :   |  ::|||  :| || ::| |||| :: |
Db     141 GKDTTKVQFNYYTNGVG---NHEKIVNLGFDAANSYHTYAFDWQPNSIKWYVDGQ-LKHT 196

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
             |:    |   :  |||:      |:| ::    || : ::||:  |
Db     197 ATTQIPQTPGK--IMMNLWNGAGVDEWLGSYN-GVTPLSRSLHWVRYTK 242


RESULT 10
S11927
licheninase (EC 3.2.1.73) precursor [validated] - Bacillus macerans
N;Alternate names: endo-beta-1,3-1,4-glucanase; lichenase
C;Species: Bacillus macerans
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2000
C;Accession: S11927
R;Borriss, R.; Buettner, K.; Maentsaelae, P.
Mol. Gen. Genet. 222, 278-283, 1990
A;Title: Structure of the beta-1,3-1,4-glucanase gene of Bacillus macerans: homologie
A;Reference number: S11927; MUID:91109712; PMID:2274030
A;Accession: S11927
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <BOR>
A;Cross-references: EMBL:X55959; NID:g296715; PIDN:CAA39426.1; PID:g296716
R;Hahn, M.; Heinemann, U.
submitted to the Brookhaven Protein Data Bank, December 1994
A;Reference number: A67074; PDB:1MAC
A;Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 26-37,'P',39-2
A;Note: recombinant form with residues 26-41 from Bacillus amyloliquefaciens expresse
R;Keitel, T.; Simon, O.; Borriss, R.; Heinemann, U.
Proc. Natl. Acad. Sci. U.S.A. 90, 5287-5291, 1993
A;Title: Molecular and active-site structure of a Bacillus 1,3-1,4-beta-glucanase.
A;Reference number: A47562; MUID:93281743; PMID:8099449
A;Contents: annotation; X-ray crystallography, 2.3 angstroms
A;Note: recombinant form with residues 26-41 from Bacillus amyloliquefaciens expresse
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin an
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-25/Domain: signal sequence #status predicted <PRO>
F;26-237/Product: licheninase #status experimental <MAT>
F;55-84/Disulfide bonds: #status experimental
F;128/Active site: Glu #status predicted

  Query Match             15.7%;  Score 226;  DB 1;  Length 237;
  Best Local Similarity   35.1%;  Pred. No. 1.5e-11;
  Matches   59;  Conservative   23;  Mismatches   70;  Indels   16;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |  ||  :     || :|  || |  :| ||| | |       | |  | |:||| |||:
Db      82 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db     139 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 191

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
            :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     192 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 235


RESULT 11
S19012
licheninase (EC 3.2.1.73) precursor - Bacillus polymyxa
N;Alternate names: endo-beta-1,3-1,4-glucanase; lichenase
C;Species: Bacillus polymyxa
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 18-Jun-1999
C;Accession: S19012
R;Gosalbes, M.J.; Perez-Gonzalez, J.A.; Gonzalez, R.; Navarro, A.
J. Bacteriol. 173, 7705-7710, 1991
A;Title: Two beta-glycanase genes are clustered in Bacillus polymyxa: molecular cloni
A;Reference number: S19011; MUID:92041687; PMID:1938968
A;Accession: S19012
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-238 <GOS>
A;Cross-references: EMBL:X57094; NID:g48815; PIDN:CAA40379.1; PID:g48817
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin an
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-26/Domain: signal sequence #status predicted <PRO>
F;27-237/Product: licheninase #status predicted <MAT>
F;56-85/Disulfide bonds: #status predicted
F;129/Active site: Glu #status predicted
```

```
  Query Match             15.6%;  Score 225;  DB 1;  Length 238;
  Best Local Similarity   34.5%;  Pred. No. 1.8e-11;
  Matches   58;  Conservative   23;  Mismatches   71;  Indels   16;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :     || :|  || |  :| ||| | |   |    |  | |:||| |||:
Db      83 FDCGEYRSTNNYGYGLYEVSMKPAKNTGIVSSFFTYTGPSH---GTQWDEIDIEFLGKDT 139

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |     ||  :   |  :|||| :| | |::| |||  :: |
Db     140 TKVQFNYYTNGVGGH---EKIINLGFDASTSFHTYAFDWQPGYIKWYVDG-VLKHT---A 192

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
           :|:  | | :  |||:     :|:| :: :  ||:   :||   |||
Db     193 TTNIPSTPGKIMMNLWNGTGVDSWLGSYNGAN-PLYAEYDWV---KYT 236


RESULT 12
JS0611
endo-1,3(4)-beta-glucanase (EC 3.2.1.6) precursor - Clostridium thermocellum
N;Alternate names: laminarinase
C;Species: Clostridium thermocellum
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jan-1997
C;Accession: JS0611; S18726
R;Zverlov, V.V.; Laptev, D.A.; Tishkov, V.I.; Velikodvorskaja, G.A.
Biochem. Biophys. Res. Commun. 181, 507-512, 1991
A;Title: Nucleotide sequence of the Clostridium thermocellum laminarinase gene.
A;Reference number: JS0611; MUID:92095946; PMID:1755832
A;Accession: JS0611
A;Molecule type: DNA
A;Residues: 1-242 <ZVE1>
A;Cross-references: EMBL:X58392
R;Zverlov, V.V.; Velikodvorskaya, G.A.
Biotechnol. Lett. 12, 811-816, 1990
A;Title: Cloning the Clostridium thermocellum thermostable laminarinase gene in Escheric
A;Reference number: S18726
A;Accession: S18726
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-242 <ZVE2>
A;Cross-references: EMBL:X58392
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991
C;Genetics:
A;Gene: lam1
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-242/Product: endo-1,3(4)-beta-glucanase #status predicted <END>

  Query Match             15.5%;  Score 222.5;  DB 2;  Length 242;
  Best Local Similarity   35.6%;  Pred. No. 3e-11;
  Matches   58;  Conservative   22;  Mismatches   64;  Indels   19;  Gaps    7;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           ::: |:  |   | :       || :| ||| |   | ||| | |   |   |  || |:|
Db      80 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 136

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || |||:    | |      |    :|  | :   | | ||||| || |:|: : |||::
Db     137 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 193

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPL 159
           | :       |:  | | :  |||       |:|::|  : ||
Db     194 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPL 231


RESULT 13
D97245
endo-1,3(4)-beta-glucanase family 16 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: D97245
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97245
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-246 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80751.1; PID:g15025848; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2807
C;Superfamily: licheninase

  Query Match             15.4%;  Score 221;  DB 2;  Length 246;
  Best Local Similarity   33.7%;  Pred. No. 4.1e-11;
  Matches   58;  Conservative   19;  Mismatches   73;  Indels   22;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           ::| |  :     || :   || |   |  || | |   |   |  || |:||| |||:
Db      91 YAGGEYRSNNRYGYGLYRVSMKPAKHIGVDSSFFSYTGPS---DNNPWDEIDIEFLGKDT 147

Qy      67 GSFQSNIITGKAGAQKTSEKH---HAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
             | |  |   |      ||   : :   | : |||||  |  ||: | |||:|| :
Db     148 TEVQFNYYTNGVG------KHEYLYKLGFDASKGFHTYGYIWEQNYIAWLVDGKEVYRA- 200

Qy     124 GGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             ||:    | :  |||      :|:| :|    |:  : ||     | |
Db     201 ---TSNIPTHPGKVMMNLWPGIGVDSWLGAYD-GVTPVKAYYNWA---MYNP 245


RESULT 14
AB3055
endo-1,3-1,4-beta-glycanase exoK [imported] - Agrobacterium tumefaciens (strain C58,
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C;Accession: AB3055
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB3055
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44856.1; PID:g17742502; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: exoK
A;Map position: linear chromosome
C;Superfamily: licheninase

  Query Match             14.3%;  Score 206;  DB 2;  Length 263;
  Best Local Similarity   35.1%;  Pred. No. 8e-10;
  Matches   54;  Conservative   26;  Mismatches   56;  Indels   18;  Gaps    6;

Qy       5 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
           ::|:  |: |    :|| :|||| |  ||  |: | |   :   | :|  |:| |||||
Db      89 RNFACGEIQTKGRYRYGTYEARMKAATGSGLNSAFFTYIGPT---DKKPHDEIDFEVLGK 145

Qy      65 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
           | |  | |  |  | |    :||   |   ||  |: |   | |  :|: |:|: |
Db     146 NTGKVQLNQYIAAKGG----NEKLVPVEGGADAGFNDYAFVWEPQRLRYYVNGKLVH--- 198

Qy     124 GGQVSNLT----GTQGLRFNLWSSESAA-WVGQF 152
             :|:: |      | : |:|| :::   |:|:|
Db     199 --EVTDETK1PQNAQKIFFSLWGTDTLKDWMGKF 230
```

```
RESULT 15
C98231
endo-1,3-1,4-beta-glycanase exoK precursor (succinoglycan biosynthesis protein exok) [im
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: C98231
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; PMID:11743194
A;Accession: C98231
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-289 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89373.1; PID:g15159224; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_1600
A;Map position: linear chromosome
C;Superfamily: licheninase

  Query Match             14.3%;  Score 206;  DB 2;  Length 289;
  Best Local Similarity   35.1%;  Pred. No. 9e-10;
  Matches   54;  Conservative   26;  Mismatches   56;  Indels   18;  Gaps    6;

Qy       5 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
           ::|:  |: |    :|| :||||| |  ||  |: | |   :   | :|  |:| |||||
Db     115 RNFACGEIQTKGRYRYGTYEARMKAATGSGLNSAFFTYIGPT---DKKPHDEIDFEVLGK 171

Qy      65 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
           | |  | |  |  | |    :||   |   ||  |: |   | |  :|: |:|: |
Db     172 NTGKVQLNQYIAAKGG----NEKLVPVEGGADAGFNDYAFVWEPQRLRYYVNGKLVH--- 224

Qy     124 GGQVSNLT----GTQGLRFNLWSSESAA-WVGQF 152
             :|:: |      | : |:|| :::   |:|:|
Db     225 --EVTDETKIPQNAQKIFFSLWGTDTLKDWMGKF 256


Search completed: January  9, 2003, 12:18:59
Job time : 13.6701 secs
```

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:14:27 ; Search time 11.7685 Seconds
(without alignments)
2025.857 Million cell updates/sec

Title: US-09-654-652A-1
Perfect score: 1333
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....TRKGQESFNGQVPRDDEPAP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1328	99.6	349	2	A44507	licheninase (EC 3.
2	266.5	20.0	851	2	H84053	endo-beta-1,3-1,4
3	265.5	19.9	252	2	A48378	licheninase (EC 3.
4	247.5	18.6	276	2	I40453	licheninase (EC 3.
5	245.5	18.4	242	1	LXBS	licheninase (EC 3.
6	242.5	18.2	802	2	A36910	xylanase, beta(1,3
7	238.5	17.9	239	1	A29091	licheninase (EC 3.
8	237	17.8	334	1	S23498	licheninase (EC 3.
9	229.5	17.2	243	1	S15388	licheninase (EC 3.
10	226	17.0	237	1	S11927	licheninase (EC 3.
11	225	16.9	238	1	S19012	licheninase (EC 3.
12	222.5	16.7	242	2	JS0611	endo-1,3(4)-beta-g
13	221	16.6	246	2	D97245	endo-1,3(4)-beta-g
14	206	15.5	263	2	AB3055	endo-1,3-1,4-beta-
15	206	15.5	289	2	C98231	endo-1,3-1,4-beta-
16	197	14.8	269	2	H95976	endo-beta-1,3-1,4-
17	190.5	14.3	302	2	G84053	hybrid-endo-beta-1
18	182.5	13.7	507	2	S64507	probable membrane
19	180	13.5	268	2	S34804	exoK protein - Rhi
20	172	12.9	282	2	T02354	xyloglucan endo-1,
21	171.5	12.9	277	2	B85354	hypothetical prote
22	169.5	12.7	282	2	A85354	hypothetical prote
23	169	12.7	642	2	B72428	laminarinase - The
24	163	12.2	286	2	S71225	xyloglucan endo-1,
25	159	11.9	287	2	T04236	xyloglucan endo-1,
26	154.5	11.6	277	2	S71222	xyloglucan endo-1,
27	150.5	11.3	269	2	S61555	xyloglucan endo-1,
28	147	11.0	467	2	S30839	UTR2 protein - yea
29	146.5	11.0	284	2	T52097	xyloglucan endo-1,
30	141.5	10.6	289	2	T06166	xyloglucan endotra
31	140.5	10.5	280	2	T02090	xyloglucan endo-1,
32	140.5	10.5	310	2	A86239	protein T10024.17
33	140	10.5	305	2	G84568	probable xylogluca
34	139.5	10.5	286	2	T06202	xyloglucan endo-1,
35	138	10.4	422	2	S48564	probable membrane
36	133.5	10.0	286	2	S48201	licheninase (EC 3.
37	132	9.9	295	2	T10210	xyloglucan endo-1,
38	131.5	9.9	292	2	T06201	xyloglucan endo-1,
39	129.5	9.7	1324	2	T18265	endo-1,3(4)-beta-g
40	128	9.6	283	2	T07678	xyloglucan endo-1,
41	127.5	9.6	269	2	T05895	xyloglucan endo-1,
42	127	9.5	292	2	T04514	xyloglucan endo-1,
43	125.5	9.4	277	2	JE0156	end-xyloglucan tra
44	125	9.4	301	2	C87296	beta-glucanase [im
45	124	9.3	299	2	D84519	probable endoxylog

ALIGNMENTS

RESULT 1
A44507
licheninase (EC 3.2.1.73) - Fibrobacter succinogenes
C;Species: Fibrobacter succinogenes
C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 08-Oct-1999
C;Accession: A44507
R;Teather, R.M.; Erfle, J.D.
J. Bacteriol. 172, 3837-3841, 1990
A;Title: DNA sequence of a Fibrobacter succinogenes mixed-linkage beta-glucanase (1,3
A;Reference number: A44507; MUID:90299807; PMID:2193918
A;Accession: A44507
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TEA>
A;Cross-references: EMBL:M33676; NID:g148575; PIDN:AAA24896.1; PID:g148576
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 99.6%; Score 1328; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.5e-104;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84

Qy      62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144

Qy     122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204

Qy     182 FTLDWTDNFDTFDSSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 FTLDWTDNFDTFDSSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 264

Qy     242 RDDEPAP 248
          |||||||
Db     265 RDDEPAP 271
```

RESULT 2
H84053
endo-beta-1,3-1,4 glucanase (licheninase) bglS [imported] - Bacillus halodurans (stra
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: H84053
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: H84053
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-851 <STO>
A;Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB06951.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: bglS

Query Match 20.0%; Score 266.5; DB 2; Length 851;
Best Local Similarity 29.7%; Pred. No. 1.8e-14;
Matches 82; Conservative 31; Mismatches 108; Indels 55; Gaps 12;

```
Qy       4 AKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLG 63
Db      99 APPYKAGELRTNDFYQYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFLG 157

Qy      64 KNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
Db     158 KDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA- 213

Qy     124 GGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP----GQGE 177
Db     214 ---TENIPQTPQKIMMNLWPGVGVDEWTGVFDGDNTPLHADYEWV---RYTPLEALDEEE 267

Qy     178 GGSDFTLDWT----------------------DNFDTFDGSRWG-KGDWTFDGNRVDLT 213
Db     268 SGDDNEEEPVEEVEEEPAEDEEVSVRPIGSAIYETFNTFNEDIWSIAHGWT-NGQMFNAT 326

Qy     214 --DKNIYSRDGMLILALTRKGQESFNGQVPRDDEPA 247
Db     327 WYDSQITFSNGIMRFAIDKE----------DDEEA 351
```

RESULT 3
A48378
licheninase (EC 3.2.1.73) - Bacillus brevis
N;Alternate names: lichenase
C;Species: Bacillus brevis
C;Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 07-May-1999
C;Accession: A48378
R;Louw, M.E.; Reid, S.J.; Watson, T.G.
Appl. Microbiol. Biotechnol. 38, 507-513, 1993
A;Title: Characterization, cloning and sequencing of a thermostable endo-(1,3-1,4) beta-
A;Reference number: A48378; MUID:93159752; PMID:7763386
A;Accession: A48378
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-252 <LOU>
A;Note: sequence extracted from NCBI backbone (NCBIN:124672, NCBIP:124673)
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 19.9%; Score 265.5; DB 2; Length 252;
Best Local Similarity 35.8%; Pred. No. 4.8e-15;
Matches 62; Conservative 24; Mismatches 74; Indels 13; Gaps 6;

```
Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db      90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
Db     149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy     123 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
Db     206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTP 251
```

RESULT 4
I40453
licheninase (EC 3.2.1.73) - Bacillus sp.
C;Species: Bacillus sp.
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 22-Jun-1999
C;Accession: I40453; S32688
R;Tabernero, C.; Coll, P.M.; Fernandez-Abalos, J.M.; Perez, P.; Santamaria, R.I.
Appl. Environ. Microbiol. 60, 1213-1220, 1994
A;Title: Cloning and DNA sequencing of bgaA, a gene encoding an endo-beta-1,3-1,4-glu
A;Reference number: I40453; MUID:94288605; PMID:7517127
A;Accession: I40453
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-276 <RES>
A;Cross-references: EMBL:Z12151; NID:g296931; PIDN:CAA78135.1; PID:g296932
C;Genetics:
A;Gene: bga1
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase

Query Match 18.6%; Score 247.5; DB 2; Length 276;
Best Local Similarity 36.1%; Pred. No. 1.8e-13;
Matches 61; Conservative 20; Mismatches 75; Indels 13; Gaps 6;

```
Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db      96 YKAGELRTNDFYHYGLFEVSMKPAKSTGTVSSFFTY-TGPWDWENDPWDEIDIEFLGKDT 154

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
Db     155 TKIQFNYFTNGVGG---NEHYHELGFDAADDFNTYAFEWRPESIRWFVNGELVHTA---- 207

Qy     127 VSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
Db     208 TENIPQTPQKIMMNLWPGIGVDGWTGRFNGEDTPVVTQYDWV---KYTP 253
```

RESULT 5
LXBS
licheninase (EC 3.2.1.73) precursor - Bacillus subtilis
N;Alternate names: 1,3-1,4-beta-D-glucan 4-gluconohydrolase; beta-glucanase; lichenas
N;Contains: licheninase E-1; licheninase E-2
C;Species: Bacillus subtilis
C;Date: 30-Jun-1988 #sequence_revision 13-Mar-1998 #text_change 16-Jun-2000
C;Accession: I40370; B69594; A22914; A90026; A90027; JU0110
R;Wolf, M.; Geczi, A.; Simon, O.; Borriss, R.
Microbiology 141, 281-290, 1995
A;Title: Genes encoding xylan and beta-glucan hydrolysing enzymes in Bacillus subtili
A;Reference number: I40370; MUID:95219081; PMID:7704256
A;Accession: I40370
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-242 <RES>
A;Cross-references: EMBL:Z46862; NID:g599673; PIDN:CAA86922.1; PID:g599674
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: B69594
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-242 <KUN>
A;Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB15943.1; PID:g26364
A;Experimental source: strain 168

```
R;Murphy, N.; McConnell, D.J.; Cantwell, B.A.
Nucleic Acids Res. 12, 5355-5367, 1984
A;Title: The DNA sequence of the gene and genetic control sites for the excreted B. subt
A;Reference number: A93526; MUID:84272222; PMID:6087283
A;Accession: A22914
A;Molecule type: DNA
A;Residues: 1-203,'L',205-242 <MUR>
A;Cross-references: EMBL:X00754; NID:g39818; PIDN:CAA25328.1; PID:g685236
A;Experimental source: strain C120
A;Note: the authors translated the codon CAA for residue 29 as Lys and CCA for residue 8
R;Tezuka, H.; Yuuki, T.; Yabuuchi, S.
Agric. Biol. Chem. 53, 2335-2339, 1989
A;Title: Construction of a beta-glucanase hyperproducing Bacillus subtilis using the clo
A;Reference number: A90026
A;Accession: A90026
A;Molecule type: DNA
A;Residues: 1-23,'S',25-82,'S',84-242 <TEZ>
A;Cross-references: DDBJ:D00518; NID:g216243; PIDN:BAA00405.1; PID:g216244
A;Experimental source: strain Y-25, clone pLE100
R;Yuuki, T.; Tezuka, H.; Yabuuchi, S.
Agric. Biol. Chem. 53, 2341-2346, 1989
A;Title: Purification and some properties of two enzymes from a beta-glucanase hyperprod
A;Reference number: A90027
A;Contents: annotation
A;Note: source was hyperproducing strain HL-25 with gene from strain Y-25
A;Note: the amino ends of the mature forms of E-1 and E-2 are pyroglutamic acid and glut
C;Genetics:
A;Gene: bglS
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin and c
C;Superfamily: licheninase
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation; p
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-242/Product: licheninase #status predicted <MAT>
F;29/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) (partial) #status

  Query Match             18.4%;  Score 245.5;  DB 1;  Length 242;
  Best Local Similarity   35.8%;  Pred. No. 2.2e-13;
  Matches   59;  Conservative   24;  Mismatches   71;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |||:
Db      87 FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFLGKDT 143

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |   |
Db     144 TKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATNQ 199

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           :   |   :  |||:      |:| ::    ||:   :||:  |
Db     200 IPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 6
A36910
xylanase, beta(1,3-1,4)-glucanase - Ruminococcus flavefaciens
C;Species: Ruminococcus flavefaciens
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
C;Accession: A36910
R;Flint, H.J.; Martin, J.; McPherson, C.A.; Daniel, A.S.; Zhang, J.X.
J. Bacteriol. 175, 2943-2951, 1993
A;Title: A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-glucanase domai
A;Reference number: A36910; MUID:93259938; PMID:8491715
A;Accession: A36910
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-802 <FLI>
A;Cross-references: GB:S61204; NID:g385910; PIDN:AAB26620.1; PID:g385911
A;Note: sequence extracted from NCBI backbone (NCBIN:131871, NCBIP:131872)
C;Superfamily: endo-1,4-beta-xylanase homology; Thermotoga xylanase A amino-terminal rep
F;42-239/Domain: endo-1,4-beta-xylanase homology <XYL>
F;259-401/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

  Query Match             18.2%;  Score 242.5;  DB 2;  Length 802;
  Best Local Similarity   34.9%;  Pred. No. 1.8e-12;
  Matches   58;  Conservative   21;  Mismatches   74;  Indels   13;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           :|| |  |     || :|  |:     | ||| | |   |   |  || |:|||:||||
Db     634 YSGGEFRTNNFYHYGYYECSMQAMKNDGVVSSFFTYTGPS---DDNPWDEIDIEILGKNT 690

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |     || : :   : :|:|||| :| |||: | |||:|| :
Db     691 TQVQFNYYTNGQGKH---EKLYDLGFDSSEAYHTYGFDWQPNYIAWYVDGREVYRA---- 743

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           ::  | | :  | |   :   |:  |:  : ||     ||   |
Db     744 TQDIPKTPGKIMMNAWPGLTVDDWLKAFN-GRTPLTAHYQWVTYNK 788


RESULT 7
A29091
licheninase (EC 3.2.1.73) beta - Bacillus amyloliquefaciens
N;Alternate names: 1,3-1,4-beta-D-glucan 4-glucanohydrolase; beta-glucanase; lichenas
C;Species: Bacillus amyloliquefaciens
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29091
R;Hofemeister, J.; Kurtz, A.; Borriss, R.; Knowles, J.
Gene 49, 177-187, 1986
A;Title: The beta-glucanase gene from Bacillus amyloliquefaciens shows extensive homo
A;Reference number: A91564; MUID:87192007; PMID:3106158
A;Accession: A29091
A;Molecule type: DNA
A;Residues: 1-239 <HOF>
A;Cross-references: GB:M15674; NID:g143009; PIDN:AAA87323.1; PID:g143010
A;Experimental source: strain BE20/78
C;Genetics:
A;Gene: bglA
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match             17.9%;  Score 238.5;  DB 1;  Length 239;
  Best Local Similarity   34.3%;  Pred. No. 8.6e-13;
  Matches   58;  Conservative   26;  Mismatches   74;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |  |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      80 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 136

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:   | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
Db     137 GKDTTKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 192

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
             |:   |  :  |||:      |:| ::    |::   :|::  |
Db     193 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYDWMRYRK 238


RESULT 8
S23498
licheninase (EC 3.2.1.73) licB precursor - Clostridium thermocellum
N;Alternate names: beta-1,3-1,4-glucanase licB; lichenase licB
C;Species: Clostridium thermocellum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S23498; S22137
R;Schimming, S.; Schwarz, W.H.; Staudenbauer, W.L.
Eur. J. Biochem. 204, 13-19, 1992
A;Title: Structure of the Clostridium thermocellum gene licB and the encoded beta-1,3
1 cellulases.
A;Reference number: S23498; MUID:92155194; PMID:1740123
A;Accession: S23498
A;Molecule type: DNA
A;Residues: 1-334 <SCH>
A;Cross-references: EMBL:X63355; NID:g40697; PIDN:CAA44959.1; PID:g40698
C;Genetics:
A;Gene: licB
```

```
C;Superfamily: licheninase licB; Clostridium cellulase repeat homology
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-334/Product: licheninase #status predicted <MAT>
F;273-296/Domain: Clostridium cellulase repeat homology <CCR1>
F;308-331/Domain: Clostridium cellulase repeat homology <CCR2>

  Query Match             17.8%;  Score 237;  DB 1;  Length 334;
  Best Local Similarity   30.9%;  Pred. No. 1.8e-12;
  Matches   80;  Conservative   27;  Mismatches   94;  Indels   58;  Gaps   13;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           ::: |:  |   | :      || :| ||| |   | ||| | |   |   |  || |:|
Db      82 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 138

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || |||:    | |       |    :|  | :   | | ||||| || |:|: : |||::
Db     139 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 195

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVY------- 169
           | :       |:  | | :  |||       |:|::|  : ||     :|| |
Db     196 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPLQAEYEYVKYYPNGVPQD 250

Qy     170 ---------KYTP-----------GQGE-GGSDFT------LDWTDNFDTFDGSRWGKG 201
                     ||             | |    ||::      |   | |   | |
Db     251 NPTPTPTIAPSTPTNPNLPLKGDVNGDGHVNSSDYSLFKRYLLRVIDRFPVGDQS---VA 307

Qy     202 DWTFDGNRVDLTDKNIYSR 220
           |   || |:| ||  :  |
Db     308 DVNRDG-RIDSTDLTMLKR 325


RESULT 9
S15388
licheninase (EC 3.2.1.73) - Bacillus licheniformis
N;Alternate names: endo-beta-1,3-1,4-D-glucanase; lichenase
C;Species: Bacillus licheniformis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S15388
R;Lloberas, J.; Perez-Pons, J.A.; Querol, E.
Eur. J. Biochem. 197, 337-343, 1991
A;Title: Molecular cloning, expression and nucleotide sequence of the endo-beta-1,3-1,4-
A;Reference number: S15388; MUID:91224124; PMID:2026156
A;Accession: S15388
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-243 <LLO>
A;Cross-references: EMBL:X57279; NID:g39558; PIDN:CAA40547.1; PID:g39559
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match             17.2%;  Score 229.5;  DB 1;  Length 243;
  Best Local Similarity   33.7%;  Pred. No. 5e-12;
  Matches   57;  Conservative   25;  Mismatches   76;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :|  || |   | ||| | |   :   || || |:||| |
Db      84 SYNKFDCGENRSVQTYGYGLYEVNMKPAKNVGIVSSFFTYTGPT---DGTPWDEIDIEFL 140

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |    |     ||   :   |  ::|||  :| || ::| |||| :: |
Db     141 GKDTTKVQFNYYTNGVG---NHEKIVNLGFDAANSYHTYAFDWQPNSIKWYVDGQ-LKHT 196

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           |:    |   :  |||:      |:| ::    || : ::||:  |
Db     197 ATTQIPQTPGK--IMMNLWNGAGVDEWLGSYN-GVTPLSRSLHWVRYTK 242


RESULT 10
S11927
licheninase (EC 3.2.1.73) precursor [validated] - Bacillus macerans
N;Alternate names: endo-beta-1,3-1,4-glucanase; lichenase
C;Species: Bacillus macerans
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2000
C;Accession: S11927
R;Borriss, R.; Buettner, K.; Maentsaelae, P.
Mol. Gen. Genet. 222, 278-283, 1990
A;Title: Structure of the beta-1,3-1,4-glucanase gene of Bacillus macerans: homologie
A;Reference number: S11927; MUID:91109712; PMID:2274030
A;Accession: S11927
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <BOR>
A;Cross-references: EMBL:X55959; NID:g296715; PIDN:CAA39426.1; PID:g296716
R;Hahn, M.; Heinemann, U.
submitted to the Brookhaven Protein Data Bank, December 1994
A;Reference number: A67074; PDB:1MAC
A;Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 26-37,'P',39-2
A;Note: recombinant form with residues 26-41 from Bacillus amyloliquefaciens expresse
R;Keitel, T.; Simon, O.; Borriss, R.; Heinemann, U.
Proc. Natl. Acad. Sci. U.S.A. 90, 5287-5291, 1993
A;Title: Molecular and active-site structure of a Bacillus 1,3-1,4-beta-glucanase.
A;Reference number: A47562; MUID:93281743; PMID:8099449
A;Contents: annotation; X-ray crystallography, 2.3 angstroms
A;Note: recombinant form with residues 26-41 from Bacillus amyloliquefaciens expresse
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin an
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-25/Domain: signal sequence #status predicted <PRO>
F;26-237/Product: licheninase #status experimental <MAT>
F;55-84/Disulfide bonds: #status experimental
F;128/Active site: Glu #status predicted

  Query Match             17.0%;  Score 226;  DB 1;  Length 237;
  Best Local Similarity   35.1%;  Pred. No. 9.7e-12;
  Matches   59;  Conservative   23;  Mismatches   70;  Indels   16;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |  ||  :     || :|  || |  :| ||| | |       | |  | |:||| |||:
Db      82 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db     139 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 191

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
           :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     192 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 235


RESULT 11
S19012
licheninase (EC 3.2.1.73) precursor - Bacillus polymyxa
N;Alternate names: endo-beta-1,3-1,4-glucanase; lichenase
C;Species: Bacillus polymyxa
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 18-Jun-1999
C;Accession: S19012
R;Gosalbes, M.J.; Perez-Gonzalez, J.A.; Gonzalez, R.; Navarro, A.
J. Bacteriol. 173, 7705-7710, 1991
A;Title: Two beta-glycanase genes are clustered in Bacillus polymyxa: molecular cloni
A;Reference number: S19011; MUID:92041687; PMID:1938968
A;Accession: S19012
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-238 <GOS>
A;Cross-references: EMBL:X57094; NID:g48815; PIDN:CAA40379.1; PID:g48817
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin an
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-26/Domain: signal sequence #status predicted <PRO>
F;27-237/Product: licheninase #status predicted <MAT>
F;56-85/Disulfide bonds: #status predicted
F;129/Active site: Glu #status predicted
```

```
  Query Match             16.9%;  Score 225;  DB 1;  Length 238;
  Best Local Similarity   34.5%;  Pred. No. 1.2e-11;
  Matches   58;  Conservative   23;  Mismatches   71;  Indels   16;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :     || :|  || |  :| ||| | |   |    |  | |:||| |||:
Db      83 FDCGEYRSTNNYGYGLYEVSMKPAKNTGIVSSFFTYTGPSH---GTQWDEIDIEFLGKDT 139

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |     ||   :   |  :|||| :| | |::| |||  :: |
Db     140 TKVQFNYYTNGVGGH---EKIINLGFDASTSFHTYAFDWQPGYIKWYVDG-VLKHT---A 192

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
           :|:  | | :  |||:     :|:| :: :  ||:   :||   |||
Db     193 TTNIPSTPGKIMMNLWNGTGVDSWLGSYNGAN-PLYAEYDWV---KYT 236


RESULT 12
JS0611
endo-1,3(4)-beta-glucanase (EC 3.2.1.6) precursor - Clostridium thermocellum
N;Alternate names: laminarinase
C;Species: Clostridium thermocellum
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jan-1997
C;Accession: JS0611; S18726
R;Zverlov, V.V.; Laptev, D.A.; Tishkov, V.I.; Velikodvorskaja, G.A.
Biochem. Biophys. Res. Commun. 181, 507-512, 1991
A;Title: Nucleotide sequence of the Clostridium thermocellum laminarinase gene.
A;Reference number: JS0611; MUID:92095946; PMID:1755832
A;Accession: JS0611
A;Molecule type: DNA
A;Residues: 1-242 <ZVE1>
A;Cross-references: EMBL:X58392
R;Zverlov, V.V.; Velikodvorskaya, G.A.
Biotechnol. Lett. 12, 811-816, 1990
A;Title: Cloning the Clostridium thermocellum thermostable laminarinase gene in Escheric
A;Reference number: S18726
A;Accession: S18726
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-242 <ZVE2>
A;Cross-references: EMBL:X58392
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991
C;Genetics:
A;Gene: lam1
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-242/Product: endo-1,3(4)-beta-glucanase #status predicted <END>

  Query Match             16.7%;  Score 222.5;  DB 2;  Length 242;
  Best Local Similarity   35.6%;  Pred. No. 2e-11;
  Matches   58;  Conservative   22;  Mismatches   64;  Indels   19;  Gaps    7;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           ::: |:  |   | :       || :| ||| |   | ||| | |   |   |  || |:|
Db      80 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 136

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || |||:    | |      |    :|  | :   | | ||||| || |:|: : |||::
Db     137 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 193

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPL 159
           | :       |:  | | :  |||       |:|::|  : ||
Db     194 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPL 231


RESULT 13
D97245
endo-1,3(4)-beta-glucanase family 16 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: D97245
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97245
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-246 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80751.1; PID:g15025848; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2807
C;Superfamily: licheninase

  Query Match             16.6%;  Score 221;  DB 2;  Length 246;
  Best Local Similarity   33.7%;  Pred. No. 2.7e-11;
  Matches   58;  Conservative   19;  Mismatches   73;  Indels   22;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           ::| |  :     || :   || |   |  || | |   |   |  || |:||| |||:
Db      91 YAGGEYRSNNRYGYGLYRVSMKPAKHIGVDSSFFSYTGPS---DNNPWDEIDIEFLGKDT 147

Qy      67 GSFQSNIITGKAGAQKTSEKH---HAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
             | |  |   |      ||    : :   | : |||||  |  ||: | |||:|| :
Db     148 TEVQFNYYTNGVG------KHEYLYKLGFDASKGFHTYGYIWEQNYIAWLVDGKEVYRA- 200

Qy     124 GGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             ||:    | :  |||      :|:| :|    |:  : ||     | |
Db     201 ---TSNIPTHPGKVMMNLWPGIGVDSWLGAYD-GVTPVKAYYNWA---MYNP 245


RESULT 14
AB3055
endo-1,3-1,4-beta-glycanase exoK [imported] - Agrobacterium tumefaciens (strain C58,
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C;Accession: AB3055
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB3055
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44856.1; PID:g17742502; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: exoK
A;Map position: linear chromosome
C;Superfamily: licheninase

  Query Match             15.5%;  Score 206;  DB 2;  Length 263;
  Best Local Similarity   35.1%;  Pred. No. 5.4e-10;
  Matches   54;  Conservative   26;  Mismatches   56;  Indels   18;  Gaps    6;

Qy       5 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
           ::|:  |: |    :|| :|||| |  ||  |: | |   :   | :|  |:| |||||
Db      89 RNFACGEIQTKGRYRYGTYEARMKAATGSGLNSAFFTYIGPT---DKKPHDEIDFEVLGK 145

Qy      65 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
           | |  | |  |  | |    :||    |   ||  |: |   | |  :|: |:|: |
Db     146 NTGKVQLNQYIAAKGG----NEKLVPVEGGADAGFNDYAFVWEPQRLRYYVNGKLVH--- 198

Qy     124 GGQVSNLT----GTQGLRFNLWSSESAA-WVGQF 152
             :|:: |      | : |:|| :::   |:|:|
Db     199 --EVTDETKIPQNAQKIFFSLWGTDTLKDWMGKF 230
```

```
RESULT 15
C98231
endo-1,3-1,4-beta-glycanase exoK precursor (succinoglycan biosynthesis protein exok) [im
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: C98231
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; PMID:11743194
A;Accession: C98231
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-289 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89373.1; PID:g15159224; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_1600
A;Map position: linear chromosome
C;Superfamily: licheninase

  Query Match             15.5%;  Score 206;  DB 2;  Length 289;
  Best Local Similarity   35.1%;  Pred. No. 6e-10;
  Matches   54;  Conservative   26;  Mismatches   56;  Indels   18;  Gaps    6;

Qy       5 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
           ::|:  |: |    :|| :|||| |  ||  |: | |   :   | :|  |:| |||||
Db     115 RNFACGEIQTKGRYRYGTYEARMKAATGSGLNSAFFTYIGPT---DKKPHDEIDFEVLGK 171

Qy      65 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
           | |  | |  |  | |    :||   |   ||  |: |   | |  :|: |:|: |
Db     172 NTGKVQLNQYIAAKGG----NEKLVPVEGGADAGFNDYAFVWEPQRLRYYVNGKLVH--- 224

Qy     124 GGQVSNLT----GTQGLRFNLWSSESAA-WVGQF 152
             :|:: |      | : |:|| :::   |:|:|
Db     225 --EVTDETKIPQNAQKIFFSLWGTDTLKDWMGKF 256


Search completed: January  9, 2003, 12:18:58
Job time : 12.7685 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:49:36 ; Search time 13.7338 Seconds
(without alignments)
747.688 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKKTAVKSALAVAAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1210	66.3	228	3	US-09-286-690-12	Sequence 12, Appl
2	248	13.6	239	1	US-08-103-998-2	Sequence 2, Appli
3	248	13.6	242	3	US-09-286-690-8	Sequence 8, Appli
4	233.5	12.8	243	3	US-09-286-690-10	Sequence 10, Appl
5	231.5	12.7	279	3	US-09-286-690-9	Sequence 9, Appli
6	227.5	12.5	238	3	US-09-286-690-7	Sequence 7, Appli
7	224.5	12.3	242	3	US-09-286-690-11	Sequence 11, Appl
8	222.5	12.2	237	1	US-08-103-998-4	Sequence 4, Appli
9	190.5	10.4	245	3	US-09-286-690-2	Sequence 2, Appli
10	159.5	8.7	306	2	US-08-824-707-2	Sequence 2, Appli
11	145	7.9	263	4	US-09-159-106-2	Sequence 2, Appli
12	144	7.9	435	4	US-09-159-106-11	Sequence 11, Appl
13	141.5	7.8	303	4	US-09-159-106-13	Sequence 13, Appl
14	133.5	7.3	276	2	US-08-712-072C-4	Sequence 4, Appli
15	119	6.5	321	2	US-08-712-072C-3	Sequence 3, Appli
16	117.5	6.4	289	2	US-08-580-545B-4	Sequence 4, Appli
17	117.5	6.4	289	4	US-09-262-653A-4	Sequence 4, Appli
18	117.5	6.4	289	4	US-08-867-484A-2	Sequence 2, Appli
19	117	6.4	620	4	US-09-269-731-6	Sequence 6, Appli
20	116.5	6.4	654	1	US-08-392-828C-2	Sequence 2, Appli
21	116.5	6.4	654	3	US-09-330-945-2	Sequence 2, Appli
22	115	6.3	262	1	US-08-392-828C-37	Sequence 37, Appl
23	115	6.3	262	3	US-09-330-945-37	Sequence 37, Appl
24	114.5	6.3	545	4	US-09-269-731-8	Sequence 8, Appli
25	113	6.2	285	2	US-08-712-072C-5	Sequence 5, Appli
26	110.5	6.1	295	1	US-08-295-657-3	Sequence 3, Appli
27	110	6.0	287	4	US-08-640-737-8	Sequence 8, Appli
28	109.5	6.0	289	4	US-08-640-737-2	Sequence 2, Appli
29	109	6.0	382	4	US-08-213-419B-19	Sequence 19, Appl
30	108.5	5.9	367	4	US-08-213-419B-15	Sequence 15, Appl
31	108.5	5.9	989	4	US-08-213-419B-2	Sequence 2, Appli
32	108.5	5.9	989	4	US-08-213-419B-4	Sequence 4, Appli
33	108	5.9	583	4	US-09-620-412C-353	Sequence 353, App
34	107.5	5.9	666	4	US-08-961-083-2	Sequence 2, Appli
35	106.5	5.8	682	3	US-08-481-435-6	Sequence 6, Appli
36	105	5.8	422	2	US-08-712-072C-2	Sequence 2, Appli
37	104	5.7	348	4	US-09-216-295-16	Sequence 16, Appl
38	103.5	5.7	395	2	US-08-404-531B-3	Sequence 3, Appli
39	103.5	5.7	395	3	US-08-476-900A-3	Sequence 3, Appli
40	103.5	5.7	395	3	US-08-488-546A-3	Sequence 3, Appli
41	103.5	5.7	1580	4	US-08-726-320-1	Sequence 1, Appli
42	103.5	5.7	1580	4	US-09-208-716-1	Sequence 1, Appli
43	103.5	5.7	1581	4	US-08-726-320-3	Sequence 3, Appli
44	103.5	5.7	1581	4	US-09-208-716-3	Sequence 3, Appli
45	98.5	5.4	1581	2	US-08-404-531B-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-286-690-12
; Sequence 12, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Fibrobacter succinogenes
US-09-286-690-12

Query Match 66.3%; Score 1210; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.3e-102;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60

Qy      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120

Qy     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180

Qy     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTF 228
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTF 228
```

RESULT 2
US-08-103-998-2
; Sequence 2, Application US/08103998
; Patent No. 5470725
; GENERAL INFORMATION:

```
;    APPLICANT:  Borriss, Rainer
;    APPLICANT:  Hofemeister, Jurgen
;    APPLICANT:  Thomsen, Karl Kristian
;    APPLICANT:  Olsen, Ole
;    APPLICANT:  Von Wettstein, Dietrich
;    TITLE OF INVENTION:  A Thermostable (1,3-1,4)-beta-Glucanase
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  1800 Diagonal Road, P.O. Box 299
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/103,998
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/773,652
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30307/123
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-9300
;      TELEFAX:  (703) 683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  239 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-103-998-2

  Query Match             13.6%;  Score 248;  DB 1;  Length 239;
  Best Local Similarity   35.4%;  Pred. No. 6.8e-15;
  Matches   62;  Conservative   27;  Mismatches   70;  Indels   16;  Gaps    8;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db      77 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 133

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |   |     ||  ::   | : ||||  :| | |::| |||  :
Db     134 EFLGKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VL 189

Qy     143 RKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           : |     :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     190 KHT---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 237


RESULT 3
US-09-286-690-8
; Sequence 8, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-286-690-8

  Query Match             13.6%;  Score 248;  DB 3;  Length 242;
  Best Local Similarity   36.0%;  Pred. No. 6.9e-15;
  Matches   63;  Conservative   25;  Mismatches   75;  Indels   12;  Gaps    7;

Qy      20 ALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVE 79
           |||:    | |   |  :::   || :| ||| |  :| ||| | |   :   || || |
Db      78 ALTSPAYNK-FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDE 133

Qy      80 VDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 139
           :||| |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||
Db     134 IDIEFLGKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDG 190

Qy     140 QEVRKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           | :: |   |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     191 Q-LKHTATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 4
US-09-286-690-10
; Sequence 10, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Bacillus licheniformis
US-09-286-690-10

  Query Match             12.8%;  Score 233.5;  DB 3;  Length 243;
  Best Local Similarity   33.7%;  Pred. No. 1.5e-13;
  Matches   58;  Conservative   26;  Mismatches   77;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :|  || |   | ||| | |   :   || || |:||
Db      81 TSPSYNKFDCGENRSVQTYGYGLYEVNMKPAKNVGIVSSFFTYTGPT---DGTPWDEIDI 137

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |   |     ||   :   |  ::|||  :| || ::| |||| :
Db     138 EFLGKDTTKVQFNYYTNGVG---NHEKIVNLGFDAANSYHTYAFDWQPNSIKWYVDGQ-L 193

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           : |   |:    |   :  |||:      |:| ::    || : ::||:  |
Db     194 KHTATTQIPQTPGK--IMMNLWNGAGVDEWLGSYN-GVTPLSRSLHWVRYTK 242


RESULT 5
US-09-286-690-9
; Sequence 9, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
```

```
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 279
;   TYPE: PRT
;   ORGANISM: Clostridium thermocellum
US-09-286-690-9

  Query Match             12.7%;  Score 231.5;  DB 3;  Length 279;
  Best Local Similarity   35.3%;  Pred. No. 2.7e-13;
  Matches   61;  Conservative   23;  Mismatches   70;  Indels   19;  Gaps    7;

Qy      28 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 81
           ::: |:  |   | :       || :| ||| |   | ||| | |   |   |  || |:|
Db      82 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 138

Qy      82 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 141
           || |||:    | |      |    :|  | :   | | ||||| || |:|: : |||::
Db     139 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 195

Qy     142 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVY 192
           | :       |:  | | :  |||       |:|::|  : ||     :|| |
Db     196 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPLQAEYEYVKYY 243


RESULT 6
US-09-286-690-7
; Sequence 7, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Bacillus polymyxa
US-09-286-690-7

  Query Match             12.5%;  Score 227.5;  DB 3;  Length 238;
  Best Local Similarity   34.3%;  Pred. No. 5e-13;
  Matches   61;  Conservative   26;  Mismatches   74;  Indels   17;  Gaps    9;

Qy      20 ALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVE 79
           :||:  : | |   |  :     || :|  || |  :| ||| | |   |    |  | |
Db      74 SLTSPANNK-FDCGEYRSTNNYGYGLYEVSMKPAKNTGIVSSFFTYTGPSH---GTQWDE 129

Qy      80 VDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 139
           :||| |||:    | |  |   |    ||   :   |  :|||| :| | |::| |||
Db     130 IDIEFLGKDTTKVQFNYYTNGVGGH---EKIINLGFDASTSFHTYAFDWQPGYIKWYVDG 186

Qy     140 QEVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           :: |    :|:  | | :  |||:     :|:| :: :  ||:   :||   |||
Db     187 -VLKHT---ATTNIPSTPGKIMMNLWNGTGVDSWLGSYNGAN-PLYAEYDWV---KYT 236


RESULT 7
US-09-286-690-11
; Sequence 11, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Clostridium thermocellum
US-09-286-690-11

  Query Match             12.3%;  Score 224.5;  DB 3;  Length 242;
  Best Local Similarity   33.5%;  Pred. No. 9.5e-13;
  Matches   62;  Conservative   25;  Mismatches   71;  Indels   27;  Gaps    8;

Qy      14 VAAAAAALTTNVS--------AKDFSGAELYTLEEVQ------YGKFEARMKMAAASGTV 59
           |:    | | ::|         ::: |:  |   | :      || :| ||| |   | |
Db      58 VSTVLEAFTGDISNGKMILTLDREYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIV 117

Qy      60 SSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQ 119
           || | |   |   |  || |:||| |||:    | |      |    :|  | :   | | ·
Db     118 SSFFTYTGPS---DNNPWDEIDIEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQ 171

Qy     120 AFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDE 177
           |||||  || |:|: : |||::| :       |:  | | :  |||       |:|::|
Db     172 DFHTYGFEWRPDYIDFYVDGKKVYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD- 226

Qy     178 SKLPL 182
           : ||
Db     227 GRTPL 231


RESULT 8
US-08-103-998-4
; Sequence 4, Application US/08103998
; Patent No. 5470725
;  GENERAL INFORMATION:
;    APPLICANT:  Borriss, Rainer
;    APPLICANT:  Hofemeister, Jurgen
;    APPLICANT:  Thomsen, Karl Kristian
;    APPLICANT:  Olsen, Ole
;    APPLICANT:  Von Wettstein, Dietrich
;    TITLE OF INVENTION:  A Thermostable (1,3-1,4)-beta-Glucanase
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  1800 Diagonal Road, P.O. Box 299
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/103,998
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/773,652
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30307/123
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-9300
;      TELEFAX:  (703) 683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  237 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-103-998-4

  Query Match             12.2%;  Score 222.5;  DB 1;  Length 237;
  Best Local Similarity   33.1%;  Pred. No. 1.4e-12;
  Matches   57;  Conservative   25;  Mismatches   79;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: :   |  ||  :     || :|  || |  :| ||| | |       | |  | |:||
Db      75 TSSAYNKFDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDI 131

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  ||     ||   :   |  |:||| :| || ::| |||| :
Db     132 EFLGKDTTKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 187

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           : |   |:    |   :  |||:      |:| ::    |::   :|::  |
Db     188 KHTATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYDWMRYRK 236


RESULT 9
US-09-286-690-2
; Sequence 2, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 245
;   TYPE: PRT
;   ORGANISM: Orpinomyces sp. PC-2
US-09-286-690-2

  Query Match             10.4%;  Score 190.5;  DB 3;  Length 245;
  Best Local Similarity   31.6%;  Pred. No. 1.2e-09;
  Matches   54;  Conservative   17;  Mismatches   87;  Indels   13;  Gaps    6;

Qy      21 LTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEV 80
           || :     ::  |  |     || |:  ||     | ||| | |   |   ||  | |:
Db      79 LTIDRDGSGYTCGEYRTKNYYGYGMFQVNMKPIKNPGVVSSFFTYTGPS---DGTKWDEI 135

Qy      81 DIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQ 140
           |||  ||  :    | |  |   |     |  | :   | | |||||  |  | : | |||
Db     136 DIEFLGYDTTKVQFNYYTNGQGHH---EHIHYLGFDASQGFHTYGFFWARNSITWYVDGT 192

Qy     141 EVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWV 189
           |          |:  | | :  | |:      |:  |:  :  :  : :||
Db     193 AVYTA----YDNIPDTPGKIMMNAWNGIGVDDWLRPFN-GRTNISAYYDWV 238


RESULT 10
US-08-824-707-2
; Sequence 2, Application US/08824707
; Patent No. 5919688
;  GENERAL INFORMATION:
;    APPLICANT:  Ferrer, Pau
;    APPLICANT:  Diers, Ivan
;    APPLICANT:  Hedegaard, Lisbeth
;    APPLICANT:  Halkier, Torben
;    APPLICANT:  Asenjo, Juan
;    APPLICANT:  Savva, Demitris
;    TITLE OF INVENTION:  No. 5919688el enzyme with beta-1,3-glucanase activity
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5919688o No. 5919688disk of No. 5919688th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/824,707
;      FILING DATE:  14-April-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gregg, Valeta A.
;      REGISTRATION NUMBER:  35,127
;      REFERENCE/DOCKET NUMBER:  4290.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  306 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-824-707-2

  Query Match             8.7%;  Score 159.5;  DB 2;  Length 306;
  Best Local Similarity   27.6%;  Pred. No. 1.1e-06;
  Matches   59;  Conservative   29;  Mismatches   89;  Indels   37;  Gaps   11;

Qy       2 NIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEV--QYGKFEARMKMAAASGTV 59
           |   : | |||              |   :: | | |   |  |:|: |||:::    |
Db     103 NYTTSRVNSALDGQGNLVITALQESDGSYTSARLTTQGNVQPQFGRIEARIQIPRGQGIW 162

Qy      60 SSMFLYQNGSEIADGRPWV---EVDI-EVLGKNP----GSFQSNIITGKAGAQKTSEKHH 111
           |: ::   |: : |  ||    |:|| | :|  |     |:      :|  |   | :
Db     163 SAFWMV--GANLPD-TPWPTSGEIDIMENVGNAPHEVHGTVHGPGYSGDNGIMGTYQHPQ 219

Qy     112 AVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAW 171
             | |  |||:|::|||  : | |||||  :     |                 : |
Db     220 GWSFADD--FHTFGIDWTPGEITWLVDGQEYHRVTTADVG---------------ANQW 261

Qy     172 VGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDF 205
           |  ||:  | |  :| | :    ||  :  : |
Db     262 V--FDQ---PFFLILN-VAIGGQWPGNPDATTPF 289
```

```
RESULT 11
US-09-159-106-2
; Sequence 2, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
;  APPLICANT: Ferrer, Pau
;  APPLICANT: Diers, Ivan
;  APPLICANT: Halkier, Torben
;  APPLICANT: Hedegaard, Lisbeth
;  TITLE OF INVENTION: An Enzyme With  -1,3-Glucanase
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 4693.204-US
;  CURRENT APPLICATION NUMBER: US/09/159,106
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: 0427/96
;  EARLIER FILING DATE: 1996-12-04
;  EARLIER APPLICATION NUMBER: 0885/96
;  EARLIER FILING DATE: 1996-08-23
;  EARLIER APPLICATION NUMBER: PCT/DK97/00160
;  EARLIER FILING DATE: 1997-04-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 263
;   TYPE: PRT
;   ORGANISM: Oerskovia xanthineolytica
US-09-159-106-2

  Query Match             7.9%;  Score 145;  DB 4;  Length 263;
  Best Local Similarity   25.9%;  Pred. No. 1.9e-05;
  Matches   49;  Conservative   29;  Mismatches   73;  Indels   38;  Gaps   10;

Qy      42 QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITG-- 99
           |||: |||:::    | :   |    ||    | ||   :|::: :| | |: : : |
Db      82 QYGRIEARIQIPRGQG-IWPAFWMLGGS--FPGTPWPSGEIDIM-ENVG-FEPHRVHGTV 136

Qy     100 -----KAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLT 154
                 |:  |    |    :    |||: ::| |  : | ||||:  :     |
Db     137 HGPGYSGGSGITGMYQHPQGWSFADTFHTFAVDWKPGEITWFVDGQQFHRVTRASVG--- 193

Qy     155 GTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS----DFTLDWT 210
                     : |||  ||:   | |  :| | |    ||  :| :       :|:
Db     194 ------------ANAWV--FDQ---PFFLILN-VAVGGQWPGYPDGTTQLPQQMKVDYV 234

Qy     211 DNFDTFDGS 219
           :|   ||
Db     235 RVYDNGSGS 243


RESULT 12
US-09-159-106-11
; Sequence 11, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
;  APPLICANT: Ferrer, Pau
;  APPLICANT: Diers, Ivan
;  APPLICANT: Halkier, Torben
;  APPLICANT: Hedegaard, Lisbeth
;  TITLE OF INVENTION: An Enzyme With  -1,3-Glucanase
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 4693.204-US
;  CURRENT APPLICATION NUMBER: US/09/159,106
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: 0427/96
;  EARLIER FILING DATE: 1996-12-04
;  EARLIER APPLICATION NUMBER: 0885/96
;  EARLIER FILING DATE: 1996-08-23
;  EARLIER APPLICATION NUMBER: PCT/DK97/00160
;  EARLIER FILING DATE: 1997-04-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
;   LENGTH: 435
;   TYPE: PRT
;   ORGANISM: Oerskovia xanthineolytica
US-09-159-106-11

  Query Match             7.9%;  Score 144;  DB 4;  Length 435;
  Best Local Similarity   21.8%;  Pred. No. 4.9e-05;
  Matches   78;  Conservative   38;  Mismatches  106;  Indels  136;  Gaps   18;

Qy       6 TAVKSALAVAAAAAALT---TNVSA--------KDFSGA---------------------- 33
             :|: :||  |||| ':|   |: :|         :| ||
Db      28 SALVAALTAAAAALAVTVAATSAAAAPGDLLWSDEFDGAAGSAPNPAVWNHETGAHGWGN 87

Qy      34 ---ELYTLE-----------------------------------EVQYGKFEARMKMAAA 55
              : ||                                    : |||: |||:::
Db      88 AELQNYTASRANSALDGQGNLVITARREGDGSYTSARMTTQGKYQPQYGRIEARIQIPRG 147

Qy      56 SGTVSSMFLYQNGSEIADGRPWV---EVDI-EVLGKNP----GSFQSNIITGKAGAQKTS 107
             | :   |    ||    | ||    |:|| | :|  |    |:      :| :|   |
Db     148 QG-IWPAFWMLGGS--FPGTPWPSSGEIDIMENVGFEPHRVHGTVHGPGYSGGSGI--TG 202

Qy     108 EKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSE 167
              |    :    |||: ::| |  : | ||||:  :     |           | |  :
Db     203 MYQHPQGWSFADTFHTFAVDWKPGEITWFVDGQQFHRVTRASVG---------ANAWVFD 253

Qy     168 ----------SAAWVGQFD-ESKLPLFQFINWVKVYKYTPGQGEGGSDFT---------- 206
                      | |  |  ::||    :::|:||    |    |:  |
Db     254 QPFFLILNVAVGGQWPGYPDGTTQLPQQMKVDYVRVYDNGSGSSSPGNPGTGLPTGTGAV 313

Qy     207 ---------LDWTDNFD-------TFDGS---RWGKGDWTFDGNRVDLTDKNIYSRDG 245
                   : | |  |       | |:    | :|    || |    | :  |||
Db     314 RAANGMCVDVPWADPTDGNPVQIVTCSGNAAQTWTRGS---DGT-VRALGKCLDVRDG 367


RESULT 13
US-09-159-106-13
; Sequence 13, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
;  APPLICANT: Ferrer, Pau
;  APPLICANT: Diers, Ivan
;  APPLICANT: Halkier, Torben
;  APPLICANT: Hedegaard, Lisbeth
;  TITLE OF INVENTION: An Enzyme With  -1,3-Glucanase
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 4693.204-US
;  CURRENT APPLICATION NUMBER: US/09/159,106
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: 0427/96
;  EARLIER FILING DATE: 1996-12-04
;  EARLIER APPLICATION NUMBER: 0885/96
;  EARLIER FILING DATE: 1996-08-23
;  EARLIER APPLICATION NUMBER: PCT/DK97/00160
;  EARLIER FILING DATE: 1997-04-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 303
;   TYPE: PRT
;   ORGANISM: Oerskovia xanthineolytica
US-09-159-106-13

  Query Match             7.8%;  Score 141.5;  DB 4;  Length 303;
  Best Local Similarity   23.0%;  Pred. No. 4.9e-05;
  Matches   68;  Conservative   34;  Mismatches   85;  Indels  109;  Gaps   15;

Qy       6 TAVKSALAVAAAAAALT---TNVSA--------KDFSGA---------------------- 33
             :|: :||  |||| |:|   |: :|         :| ||
Db      28 SALVAALTAAAAALAVTVAATSAAAAPGDLLWSDEFDGAAGSAPNPAVWNHETGAHGWGN 87

Qy      34 ---ELYTLE-----------------------------------EVQYGKFEARMKMAAA 55
              : ||                                    : |||: |||:::
Db      88 AELQNYTASRANSALDGQGNLVITARREGDGSYTSARMTTQGKYQPQYGRIEARIQIPRG 147
```

```
Qy      56 SGTVSSMFLYQNGSEIADGRPWV---EVDI-EVLGKNP----GSFQSNIITGKAGAQKTS 107
           | :   |    ||    | ||    |:|| | :|  |    |:      :| :|   |
Db     148 QG-IWPAFWMLGGS--FPGTPWPSSGEIDIMENVGFEPHRVHGTVHGPGYSGGSGI--TG 202

Qy     108 EKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSE 167
           |    :    |||: ::| |  : | ||||:  :      |
Db     203 MYQHPQGWSFADTFHTFAVDWKPGEITWFVDGQQFHRVTRASVG---------------- 246

Qy     168 SAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS----DFTLDWTDNFDTFDGS 219
           : |||  ||:   | |  :| | |    ||  :| :       :|:   :|   ||
Db     247 ANAWV--FDQ---PFFLILN-VAVGGQWPGYPDGTTQLPQQMKVDYVRVYDNGSGS 296


RESULT 14
US-08-712-072C-4
; Sequence 4, Application US/08712072C
; Patent No. 5925541
;  GENERAL INFORMATION:
;    APPLICANT:  Jack Goldstein, Alex Zhu and Lin Leng
;    TITLE OF INVENTION:  ENDO-BETA-GALACTOSIDASE
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amster, Rothstein & Ebenstein
;      STREET:  90 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH 1.44 Mb STORAGE DISKETTE
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/712,072C
;      FILING DATE:  11-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Elizabeth A.
;      REGISTRATION NUMBER:  39,911
;      REFERENCE/DOCKET NUMBER:  63475/97
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-5995
;      TELEFAX:  (212) 286-0854 or 286-0082
;      TELEX:  TWX 710-581-4766
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  gub, Rhodothermus marinus
US-08-712-072C-4

  Query Match             7.3%;  Score 133.5;  DB 2;  Length 276;
  Best Local Similarity   23.7%;  Pred. No. 0.00023;
  Matches   44;  Conservative   34;  Mismatches   85;  Indels   23;  Gaps    8;

Qy      28 KDFSGAELYTLEEVQ--YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI 82
           :::: | | |  :    ||:|| | :: :  ||  ::::  :          |    |:||
Db      93 REYTSARLVTRGKASWTYGRFEIRARLPSGRGTWPAIWMLPDRQTYGSAY-WPDNGEIDI 151

Qy      83 -EVLGKNP----GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTV 137
            | :| ||    |:  :       | |:         | |   || | :|||| :|| |
Db     152 MEHVGFNPDVVHGTVHTKAYNHLLGTQRGGSIR---VPTARTDFHVYAIEWTPEEIRWFV 208

Qy     138 DGQEVRKTEGGQVSNLTG-------TQGLRFNLWSSESAAWVGQ--FDESKLPLFQFINW 188
           |      :    ::::           |      : :   || ||   |    |    :::
Db     209 DDSLYYRFPNERLTDPEADWRHWPFDQPFHLIMNIAVGGAWGGQQGVDPEAFPAQLVVDY 268

Qy     189 VKVYKY 194
           |:||::
Db     269 VRVYRW 274


RESULT 15
US-08-712-072C-3
; Sequence 3, Application US/08712072C
; Patent No. 5925541
;  GENERAL INFORMATION:
;    APPLICANT:  Jack Goldstein, Alex Zhu and Lin Leng
;    TITLE OF INVENTION:  ENDO-BETA-GALACTOSIDASE
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amster, Rothstein & Ebenstein
;      STREET:  90 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH 1.44 Mb STORAGE DISKETTE
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/712,072C
;      FILING DATE:  11-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Elizabeth A.
;      REGISTRATION NUMBER:  39,911
;      REFERENCE/DOCKET NUMBER:  63475/97
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-5995
;      TELEFAX:  (212) 286-0854 or 286-0082
;      TELEX:  TWX 710-581-4766
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  321 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  e13b, Bacillus circulars
US-08-712-072C-3

  Query Match             6.5%;  Score 119;  DB 2;  Length 321;
  Best Local Similarity   20.4%;  Pred. No. 0.006;
  Matches   43;  Conservative   34;  Mismatches   84;  Indels   50;  Gaps    7;

Qy      20 ALTTNVSAKD------FSGAELYTLEE--VQYGKFEARMKMAAASGTVSSMFLYQNGSEI 71
           |:  : | :|      :|  :: | ::  ::||: : | |:    |   ::::    |
Db     124 AMNDSKSPQDPNRYAQYSSGKINTKDKLSLKYGRVDFRAKLPTGDGVWPALWMLPKDSVY 183

Qy      72 ADGRPWVEVDI-EVLGKNPGSFQSNIITG---KAGAQKTSEKHHAVSPAADQAFHTYGLE 127
                 |:|: |  |: |||    |  |           : |          :| | :
Db     184 GTWAASGEIDVMEARGRLPGSVSGTIHFGGQWPVNQSSGGDYHFPEGQTFANDYHVYSVV 243

Qy     128 WTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESA---------------- 169
           |  : ::| |||:   |    |              | | :|
```

```
Db    244 WEEDNIKWYVDGKFFYKVTNQQ--------------WYSTAAPNNPNAPFDEPFYLIMNL 289

Qy    170 AWVGQFD------ESKLPLFQFINWVKVYKY 194
          |  | ||       | :|    :::|:|||:
Db    290 AVGGNFDGGRTPNASDIPATMQVDYVRVYKF 320


Search completed: January  9, 2003, 12:06:35
Job time : 15.7338 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:49:36 ; Search time 9.75926 Seconds
(without alignments)
747.688 Million cell updates/sec

Title: US-09-654-652A-1
Perfect score: 1333
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....TRKGQESFNGQVPRDDEPAP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1103	82.7	228	3	US-09-286-690-12	Sequence 12, Appl
2	245.5	18.4	242	3	US-09-286-690-8	Sequence 8, Appli
3	244	18.3	239	1	US-08-103-998-2	Sequence 2, Appli
4	231.5	17.4	279	3	US-09-286-690-9	Sequence 9, Appli
5	229.5	17.2	243	3	US-09-286-690-10	Sequence 10, Appl
6	225	16.9	238	3	US-09-286-690-7	Sequence 7, Appli
7	222.5	16.7	242	3	US-09-286-690-11	Sequence 11, Appl
8	220.5	16.5	237	1	US-08-103-998-4	Sequence 4, Appli
9	189	14.2	245	3	US-09-286-690-2	Sequence 2, Appli
10	154.5	11.6	306	2	US-08-824-707-2	Sequence 2, Appli
11	145	10.9	263	4	US-09-159-106-2	Sequence 2, Appli
12	143	10.7	435	4	US-09-159-106-11	Sequence 11, Appl
13	140.5	10.5	303	4	US-09-159-106-13	Sequence 13, Appl
14	133.5	10.0	276	2	US-08-712-072C-4	Sequence 4, Appli
15	118	8.9	321	2	US-08-712-072C-3	Sequence 3, Appli
16	117.5	8.8	289	2	US-08-580-545B-4	Sequence 4, Appli
17	117.5	8.8	289	4	US-09-262-653A-4	Sequence 4, Appli
18	117.5	8.8	289	4	US-08-867-484A-2	Sequence 2, Appli
19	117	8.8	620	4	US-09-269-731-6	Sequence 6, Appli
20	116.5	8.7	654	1	US-08-392-828C-2	Sequence 2, Appli
21	116.5	8.7	654	3	US-09-330-945-2	Sequence 2, Appli
22	115	8.6	262	1	US-08-392-828C-37	Sequence 37, Appl
23	115	8.6	262	3	US-09-330-945-37	Sequence 37, Appl
24	113	8.5	285	2	US-08-712-072C-5	Sequence 5, Appli
25	113	8.5	545	4	US-09-269-731-8	Sequence 8, Appli
26	110.5	8.3	295	1	US-08-295-657-3	Sequence 3, Appli
27	110	8.3	287	4	US-08-640-737-8	Sequence 8, Appli
28	109.5	8.2	289	4	US-08-640-737-2	Sequence 2, Appli
29	105	7.9	422	2	US-08-712-072C-2	Sequence 2, Appli
30	102.5	7.7	395	2	US-08-404-531B-3	Sequence 3, Appli
31	102.5	7.7	395	3	US-08-476-900A-3	Sequence 3, Appli
32	102.5	7.7	395	3	US-08-488-546A-3	Sequence 3, Appli
33	102.5	7.7	1580	4	US-08-726-320-1	Sequence 1, Appli
34	102.5	7.7	1580	4	US-09-208-716-1	Sequence 1, Appli
35	102.5	7.7	1581	4	US-08-726-320-3	Sequence 3, Appli
36	102.5	7.7	1581	4	US-09-208-716-3	Sequence 3, Appli
37	95.5	7.2	1581	2	US-08-404-531B-6	Sequence 6, Appli
38	95.5	7.2	1581	3	US-08-476-900A-6	Sequence 6, Appli
39	95.5	7.2	1581	3	US-08-488-546A-6	Sequence 6, Appli
40	95.5	7.2	1581	4	US-08-726-320-4	Sequence 4, Appli
41	95.5	7.2	1581	4	US-09-208-716-4	Sequence 4, Appli
42	95.5	7.2	1582	2	US-08-404-531B-9	Sequence 9, Appli
43	95.5	7.2	1582	3	US-08-476-900A-9	Sequence 9, Appli
44	95.5	7.2	1582	3	US-08-488-546A-9	Sequence 9, Appli
45	95.5	7.2	1582	4	US-08-726-320-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-286-690-12
; Sequence 12, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Fibrobacter succinogenes
US-09-286-690-12

Query Match 82.7%; Score 1103; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 3e-106;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84

Qy      62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144

Qy     122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204

Qy     182 FTLDWTDNFDTFDGSRWGKGDWTF 205
           ||||||||||||||||||||||||
Db     205 FTLDWTDNFDTFDGSRWGKGDWTF 228
```

RESULT 2
US-09-286-690-8
; Sequence 8, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:

```
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-286-690-8

  Query Match             18.4%;  Score 245.5;  DB 3;  Length 242;
  Best Local Similarity   35.8%;  Pred. No. 1.5e-17;
  Matches   59;  Conservative   24;  Mismatches   71;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |||:
Db      87 FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFLGKDT 143

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |  ||     ||   :   |  |:||| :| || ::| |||| :: |   |
Db     144 TKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATNQ 199

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           :    |   :  |||:      |:| ::    ||:   :||:  |
Db     200 IPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 3
US-08-103-998-2
; Sequence 2, Application US/08103998
; Patent No. 5470725
;  GENERAL INFORMATION:
;    APPLICANT:  Borriss, Rainer
;    APPLICANT:  Hofemeister, Jurgen
;    APPLICANT:  Thomsen, Karl Kristian
;    APPLICANT:  Olsen, Ole
;    APPLICANT:  Von Wettstein, Dietrich
;    TITLE OF INVENTION:  A Thermostable (1,3-1,4)-beta-Glucanase
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  1800 Diagonal Road, P.O. Box 299
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/103,998
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/773,652
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30307/123
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-9300
;      TELEFAX:  (703) 683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  239 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-103-998-2

  Query Match             18.3%;  Score 244;  DB 1;  Length 239;
  Best Local Similarity   35.5%;  Pred. No. 2.1e-17;
  Matches   61;  Conservative   26;  Mismatches   69;  Indels   16;  Gaps    8;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      80 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 136

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db     137 GKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT 192

Qy     123 EGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
               :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     193 ---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 237


RESULT 4
US-09-286-690-9
; Sequence 9, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 279
;   TYPE: PRT
;   ORGANISM: Clostridium thermocellum
US-09-286-690-9

  Query Match             17.4%;  Score 231.5;  DB 3;  Length 279;
  Best Local Similarity   35.3%;  Pred. No. 5.2e-16;
  Matches   61;  Conservative   23;  Mismatches   70;  Indels   19;  Gaps    7;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           ::: |:  |   | :      || :| ||| |   | ||| | |   |   |  || |:|
Db      82 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 138

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || |||:    | |      |   :|  | :   | | ||||| || |:|: : |||::
Db     139 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 195

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVY 169
           | :       |:  | | :  |||       |:|::|  : ||     :|| |
Db     196 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPLQAEYEYVKYY 243


RESULT 5
US-09-286-690-10
; Sequence 10, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
```

```
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Bacillus licheniformis
US-09-286-690-10

  Query Match             17.2%;  Score 229.5;  DB 3;  Length 243;
  Best Local Similarity   33.7%;  Pred. No. 6.9e-16;
  Matches   57;  Conservative   25;  Mismatches   76;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :|  || |   | ||| | |   :   || || |:||| |
Db      84 SYNKFDCGENRSVQTYGYGLYEVNMKPAKNVGIVSSFFTYTGPT---DGTPWDEIDIEFL 140

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |   |     ||   :   |  ::|||  :| || ::| |||| :: |
Db     141 GKDTTKVQFNYYTNGVG---NHEKIVNLGFDAANSYHTYAFDWQPNSIKWYVDGQ-LKHT 196

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
             |:    |   :  |||:      |:| ::    || : ::||:  |
Db     197 ATTQIPQTPGK--IMMNLWNGAGVDEWLGSYN-GVTPLSRSLHWVRYTK 242


RESULT 6
US-09-286-690-7
; Sequence 7, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Bacillus polymyxa
US-09-286-690-7

  Query Match             16.9%;  Score 225;  DB 3;  Length 238;
  Best Local Similarity   34.5%;  Pred. No. 1.9e-15;
  Matches   58;  Conservative   23;  Mismatches   71;  Indels   16;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :    || :|  || |  :| ||| | |   |    |  | |:||| |||:
Db      83 FDCGEYRSTNNYGYGLYEVSMKPAKNTGIVSSFFTYTGPSH---GTQWDEIDIEFLGKDT 139

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |     ||   :   |  :|||| :| | |::| |||  :: |
Db     140 TKVQFNYYTNGVGGH---EKIINLGFDASTSFHTYAFDWQPGYIKWYVDG-VLKHT---A 192

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
           :|:  | | :  |||:     :|:| :: :  ||:   :||   |||
Db     193 TTNIPSTPGKIMMNLWNGTGVDSWLGSYNGAN-PLYAEYDWV---KYT 236


RESULT 7
US-09-286-690-11
; Sequence 11, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Clostridium thermocellum
US-09-286-690-11

  Query Match             16.7%;  Score 222.5;  DB 3;  Length 242;
  Best Local Similarity   35.6%;  Pred. No. 3.6e-15;
  Matches   58;  Conservative   22;  Mismatches   64;  Indels   19;  Gaps    7;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           ::: |:  |   | :       || :| ||| |   | ||| | |   |   |  || |:|
Db      80 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 136

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || |||:    | |       |    :|  | :   | | ||||| || |:|: : |||::
Db     137 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 193

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPL 159
           | :        |:  | | :  |||       |:|::|  : ||
Db     194 VYR----GTRNIPYTPGKIMMNLWPGIGVDEWLGRYD-GRTPL 231


RESULT 8
US-08-103-998-4
; Sequence 4, Application US/08103998
; Patent No. 5470725
;  GENERAL INFORMATION:
;    APPLICANT:  Borriss, Rainer
;    APPLICANT:  Hofemeister, Jurgen
;    APPLICANT:  Thomsen, Karl Kristian
;    APPLICANT:  Olsen, Ole
;    APPLICANT:  Von Wettstein, Dietrich
;    TITLE OF INVENTION:  A Thermostable (1,3-1,4)-beta-Glucanase
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  1800 Diagonal Road, P.O. Box 299
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/103,998
;      FILING DATE:
;      CLASSIFICATION:  435
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/773,652
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30307/123
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-9300
;      TELEFAX:  (703) 683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  237 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-103-998-4

  Query Match             16.5%;  Score 220.5;  DB 1;  Length 237;
  Best Local Similarity   33.9%;  Pred. No. 5.6e-15;
  Matches   56;  Conservative   23;  Mismatches   75;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |  ||  :    || :|  || |  :| ||| | |      | |  | |:||| |||:
Db      82 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |  ||     ||    :   |  |:|||  :| || ::| |||| :: |   |
Db     139 TKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATTQ 194

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           :    |   :  |||:      |:| ::    |::   :|::  |
Db     195 IPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYDWMRYRK 236


RESULT 9
US-09-286-690-2
; Sequence 2, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
;  APPLICANT: Li, Xin-Liang
;  APPLICANT: Ljungdahl, Lars G.
;  APPLICANT: Chen, Huizhong
;  TITLE OF INVENTION: Lichenase and Coding Sequences
;  FILE REFERENCE: 55-96
;  CURRENT APPLICATION NUMBER: US/09/286,690
;  CURRENT FILING DATE: 1999-04-05
;  EARLIER APPLICATION NUMBER: US 60/027,882
;  EARLIER FILING DATE: 1999-10-04
;  EARLIER APPLICATION NUMBER: PCT/US97/17811
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 245
;   TYPE: PRT
;   ORGANISM: Orpinomyces sp. PC-2
US-09-286-690-2

  Query Match             14.2%;  Score 189;  DB 3;  Length 245;
  Best Local Similarity   32.9%;  Pred. No. 1.1e-11;
  Matches   56;  Conservative   16;  Mismatches   76;  Indels   22;  Gaps    8;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           :| ||   ||  | :      || |:  ||     | ||| | |   |   ||  | |:|
Db      83 RDGSG---YTCGEYRTKNYYGYGMFQVNMKPIKNPGVVSSFFTYTGPS---DGTKWDEID 136

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || || :    | |  |   |     |  | :   | | |||||  |  | : | |||
Db     137 IEFLGYDTTKVQFNYYTNGQGHH---EHIHYLGFDASQGFHTYGFFWARNSITWYVDGTA 193

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWV 166
           |        |:  | | :  | |:      |:  |:  :  :  : :||
Db     194 VYTA----YDNIPDTPGKIMMNAWNGIGVDDWLRPFN-GRTNISAYYDWV 238


RESULT 10
US-08-824-707-2
; Sequence 2, Application US/08824707
; Patent No. 5919688
;  GENERAL INFORMATION:
;    APPLICANT:  Ferrer, Pau
;    APPLICANT:  Diers, Ivan
;    APPLICANT:  Hedegaard, Lisbeth
;    APPLICANT:  Halkier, Torben
;    APPLICANT:  Asenjo, Juan
;    APPLICANT:  Savva, Demitris
;    TITLE OF INVENTION:  No. 5919688el enzyme with beta-1,3-glucanase activity
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5919688o No. 5919688disk of No. 5919688th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/824,707
;      FILING DATE:  14-April-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gregg, Valeta A.
;      REGISTRATION NUMBER:  35,127
;      REFERENCE/DOCKET NUMBER:  4290.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  306 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-824-707-2

  Query Match             11.6%;  Score 154.5;  DB 2;  Length 306;
  Best Local Similarity   28.5%;  Pred. No. 5.5e-08;
  Matches   53;  Conservative   28;  Mismatches   68;  Indels   37;  Gaps   11;

Qy       7 FSGAELYTLEEV--QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI-E 60
           :: | | |    |  |:|: |||:::    |  |: ::   |: : |  ||    |:|| |
Db     131 YTSARLTTQGNVQPQFGRIEARIQIPRGQGIWSAFWMV--GANLPD-TPWPTSGEIDIME 187

Qy      61 VLGKNP----GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 116
            :|  |    |:       :|  |   | :     | | |  |||:|::|||  : | |||
Db     188 NVGNAPHEVHGTVHGPGYSGDNGIMGTYQHPQGWSFADD--FHTFGIDWTPGEITWLVDG 245

Qy     117 QEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQG 176
           ||  :     |                 :  ||  ||:   | |  :| | :    ||
Db     246 QEYHRVTTADVG---------------ANQWV--FDQ---PFFLILN-VAIGGQWPGNP 283

Qy     177 EGGSDF 182
           :  : |
Db     284 DATTPF 289


RESULT 11
US-09-159-106-2
; Sequence 2, Application US/09159106
; Patent No. 6284509
```

```
; GENERAL INFORMATION:
;  APPLICANT: Ferrer, Pau
;  APPLICANT: Diers, Ivan
;  APPLICANT: Halkier, Torben
;  APPLICANT: Hedegaard, Lisbeth
;  TITLE OF INVENTION: An Enzyme With  -1,3-Glucanase
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 4693.204-US
;  CURRENT APPLICATION NUMBER: US/09/159,106
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: 0427/96
;  EARLIER FILING DATE: 1996-12-04
;  EARLIER APPLICATION NUMBER: 0885/96
;  EARLIER FILING DATE: 1996-08-23
;  EARLIER APPLICATION NUMBER: PCT/DK97/00160
;  EARLIER FILING DATE: 1997-04-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 263
;   TYPE: PRT
;   ORGANISM: Oerskovia xanthineolytica
US-09-159-106-2

  Query Match             10.9%;  Score 145;  DB 4;  Length 263;
  Best Local Similarity   25.9%;  Pred. No. 4.2e-07;
  Matches   49;  Conservative   29;  Mismatches   73;  Indels   38;  Gaps   10;

Qy      19 QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITG-- 76
           |||: |||:::    | :   |    ||    | ||   :|::: :| | |: : : |
Db      82 QYGRIEARIQIPRGQG-IWPAFWMLGGS--FPGTPWPSGEIDIM-ENVG-FEPHRVHGTV 136

Qy      77 -----KAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLT 131
                |:  |    |    :    |||: ::| |  : | ||||:  :     |
Db     137 HGPGYSGGGSGITGMYQHPQGWSFADTFHTFAVDWKPGEITWFVDGQQFHRVTRASVG--- 193

Qy     132 GTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS----DFTLDWT 187
                       : |||  ||:   | |  :| | |    ||  :| :       :|:
Db     194 -------------ANAWV--FDQ---PFFLILN-VAVGGQWPGYPDGTTQLPQQMKVDYV 234

Qy     188 DNFDTFDGS 196
           :|   ||
Db     235 RVYDNGSGS 243


RESULT 12
US-09-159-106-11
; Sequence 11, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
;  APPLICANT: Ferrer, Pau
;  APPLICANT: Diers, Ivan
;  APPLICANT: Halkier, Torben
;  APPLICANT: Hedegaard, Lisbeth
;  TITLE OF INVENTION: An Enzyme With  -1,3-Glucanase
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 4693.204-US
;  CURRENT APPLICATION NUMBER: US/09/159,106
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: 0427/96
;  EARLIER FILING DATE: 1996-12-04
;  EARLIER APPLICATION NUMBER: 0885/96
;  EARLIER FILING DATE: 1996-08-23
;  EARLIER APPLICATION NUMBER: PCT/DK97/00160
;  EARLIER FILING DATE: 1997-04-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
;   LENGTH: 435
;   TYPE: PRT
;   ORGANISM: Oerskovia xanthineolytica
US-09-159-106-11

  Query Match             10.7%;  Score 143;  DB 4;  Length 435;
  Best Local Similarity   24.6%;  Pred. No. 1.4e-06;
  Matches   62;  Conservative   29;  Mismatches   95;  Indels   66;  Gaps   14;

Qy      19 QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI-EVLGKNP----GSFQ 70
           |||: |||:::    | :   |    ||    | ||    |:|| | :|  |    |:
Db     134 QYGRIEARIQIPRGQG-IWPAFWMLGGS--FPGTPWPSSGEIDIMENVGFEPHRVHGTVH 190

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
              :| :|   |    |    :    |||: ::| |  : | ||||:  :     |
Db     191 GPGYSGGSGI--TGMYQHPQGWSFADTFHTFAVDWKPGEITWFVDGQQFHRVTRASVG-- 246

Qy     131 TGTQGLRFNLWSSE-----------SAAWVGQFD-ESKLPLFQFINWVKVYKYTPGQGEG 178
                  | |  :              | |  |  ::||    :::|:||    |
Db     247 -------ANAWVFDQPFFLILNVAVGGQWPGYPDGTTQLPQQMKVDYVRVYDNGSGSSSP 299

Qy     179 GSDFT-----------------LDWTDNFD-------TFDGS---RWGKGDWTFDGNRV 210
           |:  |                 : | |  |       |  |:    | :|    ||  |
Db     300 GNPGTGLPTGTGAVRAANGMCVDVPWADPTDGNPVQIVTCSGNAAQTWTRGS---DGT-V 355

Qy     211 DLTDKNIYSRDG 222
              | :  |||
Db     356 RALGKCLDVRDG 367


RESULT 13
US-09-159-106-13
; Sequence 13, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
;  APPLICANT: Ferrer, Pau
;  APPLICANT: Diers, Ivan
;  APPLICANT: Halkier, Torben
;  APPLICANT: Hedegaard, Lisbeth
;  TITLE OF INVENTION: An Enzyme With  -1,3-Glucanase
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 4693.204-US
;  CURRENT APPLICATION NUMBER: US/09/159,106
;  CURRENT FILING DATE: 1998-09-23
;  EARLIER APPLICATION NUMBER: 0427/96
;  EARLIER FILING DATE: 1996-12-04
;  EARLIER APPLICATION NUMBER: 0885/96
;  EARLIER FILING DATE: 1996-08-23
;  EARLIER APPLICATION NUMBER: PCT/DK97/00160
;  EARLIER FILING DATE: 1997-04-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 303
;   TYPE: PRT
;   ORGANISM: Oerskovia xanthineolytica
US-09-159-106-13

  Query Match             10.5%;  Score 140.5;  DB 4;  Length 303;
  Best Local Similarity   27.4%;  Pred. No. 1.5e-06;
  Matches   52;  Conservative   25;  Mismatches   74;  Indels   39;  Gaps   11;

Qy      19 QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI-EVLGKNP----GSFQ 70
           |||: |||:::    | :   |    ||    | ||    |:|| | :|  |    |:
Db     134 QYGRIEARIQIPRGQG-IWPAFWMLGGS--FPGTPWPSSGEIDIMENVGFEPHRVHGTVH 190

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
              :| :|   |    |    :    |||: ::| |  : | ||||:  :     |
Db     191 GPGYSGGSGI--TGMYQHPQGWSFADTFHTFAVDWKPGEITWFVDGQQFHRVTRASVG-- 246

Qy     131 TGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS----DFTLDW 186
                       : |||  ||:   | |  :| | |    ||  :| :       :|:
Db     247 -------------ANAWV--FDQ---PFFLILN-VAVGGQWPGYPDGTTQLPQQMKVDY 286

Qy     187 TDNFDTFDGS 196
              :|   ||
Db     287 VRVYDNGSGS 296
```

```
RESULT 14
US-08-712-072C-4
; Sequence 4, Application US/08712072C
; Patent No. 5925541
;  GENERAL INFORMATION:
;    APPLICANT:  Jack Goldstein, Alex Zhu and Lin Leng
;    TITLE OF INVENTION:  ENDO-BETA-GALACTOSIDASE
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amster, Rothstein & Ebenstein
;      STREET:  90 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH 1.44 Mb STORAGE DISKETTE
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/712,072C
;      FILING DATE:  11-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Elizabeth A.
;      REGISTRATION NUMBER:  39,911
;      REFERENCE/DOCKET NUMBER:  63475/97
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-5995
;      TELEFAX:  (212) 286-0854 or 286-0082
;      TELEX:  TWX 710-581-4766
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  gub, Rhodothermus marinus
US-08-712-072C-4

  Query Match             10.0%;  Score 133.5;  DB 2;  Length 276;
  Best Local Similarity   23.7%;  Pred. No. 7e-06;
  Matches   44;  Conservative   34;  Mismatches   85;  Indels   23;  Gaps    8;

Qy       5 KDFSGAELYTLEEVQ--YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI 59
           :::: | | |   :    ||:|| | :: :  ||  ::::  :         |    |:||
Db      93 REYTSARLVTRGKASWTYGRFEIRARLPSGRGTWPAIWMLPDRQTYGSAY-WPDNGEIDI 151

Qy      60 -EVLGKNP----GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTV 114
            | :| ||    |:  :       | |:         | |   || | :|||| :|| |
Db     152 MEHVGFNPDVVHGTVHTKAYNHLLGTQRGGSIR---VPTARTDFHVYAIEWTPEEIRWFV 208

Qy     115 DGQEVRKTEGGQVSNLTG-------TQGLRFNLWSSESAAWVGQ--FDESKLPLFQFINW 165
           |    :    ::::          |     :  :   || ||   |    |    :::
Db     209 DDSLYYRFPNERLTDPEADWRHWPFDQPFHLIMNIAVGGAWGGQQGVDPEAFPAQLVVDY 268

Qy     166 VKVYKY 171
           |:||::
Db     269 VRVYRW 274


RESULT 15
US-08-712-072C-3
```

```
; Sequence 3, Application US/08712072C
; Patent No. 5925541
;  GENERAL INFORMATION:
;    APPLICANT:  Jack Goldstein, Alex Zhu and Lin Leng
;    TITLE OF INVENTION:  ENDO-BETA-GALACTOSIDASE
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amster, Rothstein & Ebenstein
;      STREET:  90 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH 1.44 Mb STORAGE DISKETTE
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/712,072C
;      FILING DATE:  11-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Elizabeth A.
;      REGISTRATION NUMBER:  39,911
;      REFERENCE/DOCKET NUMBER:  63475/97
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-5995
;      TELEFAX:  (212) 286-0854 or 286-0082
;      TELEX:  TWX 710-581-4766
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  321 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  el3b, Bacillus circulars
US-08-712-072C-3

  Query Match             8.9%;  Score 118;  DB 2;  Length 321;
  Best Local Similarity   20.5%;  Pred. No. 0.00035;
  Matches   40;  Conservative   31;  Mismatches   80;  Indels   44;  Gaps    6;

Qy       7 FSGAELYTLEE--VQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI-EVLG 63
           :|  :: | ::  ::||: : | |:    |   ::::    |         |:|: |  |
Db     140 YSSGKINTKDKLSLKYGRVDFRAKLPTGDGVWPALWMLPKDSVYGTWAASGEIDVMEARG 199

Qy      64 KNPGSFQSNIITG---KAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120
           : |||    |  |           : |          :| | : |  : ::| |||:
Db     200 RLPGSVSGTIHFGGQWPVNQSSGGDYHFPEGQTFANDYHVYSVVWEEDNIKWYVDGKFFY 259

Qy     121 KTEGGQVSNLTGTQGLRFNLWSSESA----------------AWVGQFD------ESK 156
           |   |              | | :|                 |  | ||       |
Db     260 KVTNQQ-------------WYSTAAPNNPNAPFDEPFYLIMNLAVGGNFDGGRTPNASD 305

Qy     157 LPLFQFINWVKVYKY 171
           :|    :::|:|||:
Db     306 IPATMQVDYVRVYKF 320


Search completed: January  9, 2003, 12:06:32
Job time : 10.7593 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:50:42 ; Search time 164.402 Seconds
(without alignments)
1368.672 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKKTAVKSALAVAAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1824	100.0	349	20	US-09-654-652A-3	Sequence 3, Appli
2	1824	100.0	349	21	US-09-791-537-30441	Sequence 30441, A
3	1340	73.5	267	20	US-09-654-652A-2	Sequence 2, Appli
4	1328	72.8	248	20	US-09-654-652A-1	Sequence 1, Appli
5	272.5	14.9	214	21	US-09-791-537-145019	Sequence 145019,
6	266.5	14.6	259	21	US-09-791-537-106873	Sequence 106873,
7	265.5	14.6	252	21	US-09-791-537-32335	Sequence 32335, A
8	257.5	14.1	308	18	US-09-463-862A-1	Sequence 1, Appli
9	256.5	14.1	208	21	US-09-791-537-6910	Sequence 6910, Ap
10	252.5	13.8	394	21	US-09-791-537-132244	Sequence 132244,
11	250.5	13.7	242	21	US-09-791-537-27063	Sequence 27063, A
12	249	13.7	239	21	US-09-791-537-103347	Sequence 103347,
13	248	13.6	239	21	US-09-791-537-103324	Sequence 103324,
14	248	13.6	242	21	US-09-791-537-73450	Sequence 73450, A
15	248	13.6	242	21	US-09-791-537-79889	Sequence 79889, A
16	247.5	13.6	214	21	US-09-791-537-117361	Sequence 117361,
17	247.5	13.6	276	21	US-09-791-537-31771	Sequence 31771, A
18	242.5	13.3	239	21	US-09-791-537-30436	Sequence 30436, A
19	242.5	13.3	802	21	US-09-791-537-50107	Sequence 50107, A
20	242.5	13.3	802	21	US-09-791-537-116925	Sequence 116925,
21	239.5	13.1	214	21	US-09-791-537-88552	Sequence 88552, A
22	238.5	13.1	239	21	US-09-791-537-18679	Sequence 18679, A
23	237.5	13.0	453	27	US-60-385-568-289	Sequence 289, App
24	237.5	13.0	462	16	US-09-248-796-14808	Sequence 14808, A
25	237.5	13.0	462	25	US-10-179-131-9854	Sequence 9854, Ap
26	237.5	13.0	462	27	US-60-096-409-14808	Sequence 14808, A
27	237	13.0	239	21	US-09-791-537-103817	Sequence 103817,
28	237	13.0	334	21	US-09-791-537-48909	Sequence 48909, A
29	233.5	12.8	243	21	US-09-791-537-30440	Sequence 30440, A
30	231.5	12.7	320	21	US-09-791-537-67048	Sequence 67048, A
31	230	12.6	237	21	US-09-791-537-72848	Sequence 72848, A
32	230	12.6	507	27	US-60-360-039-21990	Sequence 21990, A
33	228	12.5	212	21	US-09-791-537-88873	Sequence 88873, A
34	228	12.5	214	21	US-09-791-537-76270	Sequence 76270, A
35	228	12.5	237	21	US-09-791-537-103320	Sequence 103320,
36	228	12.5	237	21	US-09-791-537-106876	Sequence 106876,
37	228	12.5	239	21	US-09-791-537-103345	Sequence 103345,
38	228	12.5	239	21	US-09-791-537-103350	Sequence 103350,
39	228	12.5	239	21	US-09-791-537-103355	Sequence 103355,
40	228	12.5	239	21	US-09-791-537-103357	Sequence 103357,
41	228	12.5	239	21	US-09-791-537-103814	Sequence 103814,
42	228	12.5	239	21	US-09-791-537-103819	Sequence 103819,
43	228	12.5	240	21	US-09-791-537-64827	Sequence 64827, A
44	227.5	12.5	238	21	US-09-791-537-79887	Sequence 79887, A
45	224.5	12.3	242	21	US-09-791-537-37095	Sequence 37095, A

ALIGNMENTS

RESULT 1
US-09-654-652A-3
; Sequence 3, Application US/09654652A
; GENERAL INFORMATION:
; APPLICANT: SHYUR, LIE-FEN
; APPLICANT: CHEN, JUI-LIN
; APPLICANT: YANG, NING-SUN
; TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1,
; TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
; TITLE OF INVENTION: THERMO-TOLERANCE
; FILE REFERENCE: 4910-8
; CURRENT APPLICATION NUMBER: US/09/654,652A
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Fibrobacter succinogenes
US-09-654-652A-3

Query Match 100.0%; Score 1824; DB 20; Length 349;
Best Local Similarity 100.0%; Pred. No: 1.3e-151;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60
```

```
Qy      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120

Qy     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180

Qy     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 240

Qy     241 YSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 YSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV 300

Qy     301 PPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNGHKRYRVNFEH 349
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 PPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNGHKRYRVNFEH 349


RESULT 2
US-09-791-537-30441
; Sequence 30441, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30441
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Fibrobacter succinogenes
US-09-791-537-30441

  Query Match             100.0%;  Score 1824;  DB 21;  Length 349;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-151;
  Matches  349;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60

Qy      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120

Qy     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180

Qy     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 240

Qy     241 YSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 YSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV 300

Qy     301 PPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNGHKRYRVNFEH 349
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 PPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNGHKRYRVNFEH 349


RESULT 3
US-09-654-652A-2
; Sequence 2, Application US/09654652A
; GENERAL INFORMATION:
;  APPLICANT: SHYUR, LIE-FEN
;  APPLICANT:  CHEN, JUI-LIN
;  APPLICANT:  YANG, NING-SUN
;  TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1,
;  TITLE OF INVENTION:  4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
;  TITLE OF INVENTION:  THERMO-TOLERANCE
;  FILE REFERENCE: 4910-8
;  CURRENT APPLICATION NUMBER: US/09/654,652A
;  CURRENT FILING DATE:  2000-09-05
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 267
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme
;   OTHER INFORMATION: with enhanced activity and thermal stability
US-09-654-652A-2

  Query Match             73.5%;  Score 1340;  DB 20;  Length 267;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-109;
  Matches  250;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61

Qy      85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121

Qy     145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181

Qy     205 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 264
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 241

Qy     265 RDDEPAPQSSS 275
           ||||||| |||
Db     242 RDDEPAPNSSS 252


RESULT 4
US-09-654-652A-1
; Sequence 1, Application US/09654652A
; GENERAL INFORMATION:
;  APPLICANT: SHYUR, LIE-FEN
;  APPLICANT:  CHEN, JUI-LIN
;  APPLICANT:  YANG, NING-SUN
;  TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1,
;  TITLE OF INVENTION:  4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
;  TITLE OF INVENTION:  THERMO-TOLERANCE
;  FILE REFERENCE: 4910-8
;  CURRENT APPLICATION NUMBER: US/09/654,652A
;  CURRENT FILING DATE:  2000-09-05
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 248
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme
;   OTHER INFORMATION: with enhanced activity and thermal stability
US-09-654-652A-1

  Query Match             72.8%;  Score 1328;  DB 20;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-108;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61

Qy      85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121

Qy     145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181

Qy     205 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 264
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 241

Qy     265 RDDEPAP 271
           |||||||
Db     242 RDDEPAP 248
```

RESULT 5
US-09-791-537-145019
; Sequence 145019, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 145019
; LENGTH: 214
; TYPE: PRT
; ORGANISM: pdb 1CPM
US-09-791-537-145019

Query Match 14.9%; Score 272.5; DB 21; Length 214;
Best Local Similarity 33.8%; Pred. No. 3.5e-15;
Matches 77; Conservative 33; Mismatches 93; Indels 25; Gaps 12;

```
Qy      30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
           |  ||  :     || :|  || |  :| ||| | |      | |  | |:||| |||:
Db       1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 57

Qy      90 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 149
             | |  |   |     ||  ::   | : |||| :| | |::| |||  :: |
Db      58 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 110

Qy     150 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTL 207
           :|:  | | :  |||:      |:| :: :  ||:   :||   |||  |  ||| |
Db     111 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYTSNQ-TGGSFF-- 163

Qy     208 DWTDNFDTFDGSRWGKGDWTFDGNRVDLT--DKNI-YSRDGMLILALT 252
             : |::::   | | |   :|   : |    |: :: || | | ||
Db     164 ---EPFNSYNSGTWEKADGYSNGGVFNCTWRANNVNFTNDGKLKLGLT 208
```

RESULT 6
US-09-791-537-106873
; Sequence 106873, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 106873
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Brevibacillus brevis
US-09-791-537-106873

Query Match 14.6%; Score 266.5; DB 21; Length 259;
Best Local Similarity 36.0%; Pred. No. 1.6e-14;
Matches 63; Conservative 24; Mismatches 75; Indels 13; Gaps 6;

```
Qy      26 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 85
           ||:::   || | :   || ||  || |   ||||| | |   |    || || |:||| |
Db      90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy      86 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 145
           ||:    | |  |   |    :| :: :   | ::|:||  ||  : : | |:|: |
Db     149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy     146 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQ 198
                |:  | | :  |||       | | ||    |:: : :||   :||| |
Db     206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTPLQ 253
```

RESULT 7
US-09-791-537-32335
; Sequence 32335, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32335
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Brevibacillus brevis
US-09-791-537-32335

Query Match 14.6%; Score 265.5; DB 21; Length 252;
Best Local Similarity 35.8%; Pred. No. 1.8e-14;
Matches 62; Conservative 24; Mismatches 74; Indels 13; Gaps 6;

```
Qy      26 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 85
           ||:::   || | :   || ||  || |   ||||| | |   |    || || |:||| |
Db      90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy      86 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 145
           ||:    | |  |   |    :| :: :   | ::|:||  ||  : : | |:|: |
Db     149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy     146 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 196
                |:  | | :  |||       | | ||    |:: : :||   :|||
Db     206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTP 251
```

RESULT 8
US-09-463-862A-1
; Sequence 1, Application US/09463862A
; GENERAL INFORMATION:
; APPLICANT: Hillen, Wolfgang
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: New Beta-Glucanase from a Bacillus
; FILE REFERENCE: H2081PCT/US
; CURRENT APPLICATION NUMBER: US/09/463,862A

```
;  CURRENT FILING DATE:  2000-05-01
;  PRIOR APPLICATION NUMBER: DE 197 32 751.6
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: PCT/EP98/04564
;  PRIOR FILING DATE: 1998-07-21
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 308
;   TYPE: PRT
;   ORGANISM: Bacillus alkalophilus DSM 9956
US-09-463-862A-1

  Query Match             14.1%;  Score 257.5;  DB 18;  Length 308;
  Best Local Similarity   36.5%;  Pred. No. 1.3e-13;
  Matches   62;  Conservative   22;  Mismatches   73;  Indels   13;  Gaps    6;

Qy      29 DFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN 88
           ::   || | :  ||| ||  || | ::|||||:| |  |    |  || |:||| |||:
Db      96 EYKAGELRTNQFYQYGLFEVNMKPAKSTGTVSSLFTY-TGPWDWDNDPWDEIDIEFLGKD 154

Qy      89 PGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGG 148
             | |  |   |    :| :| :   | ::|:||  || |  : | |:|: |
Db     155 TTRVQFNYFTNGVG---NNEHYHELGFDASESFNTYAFEWRPESISWYVNGELVYTA--- 208

Qy     149 QVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 196
             |:  | | :  |||       | | ||    |:    :||   :|||
Db     209 -TENIPQTPQKIMMNLWPGIGVDGWTGVFDGEDTPVVTEYDWV---RYTP 254


RESULT 9
US-09-791-537-6910
; Sequence 6910, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6910
;   LENGTH: 208
;   TYPE: PRT
;   ORGANISM: pdb 1CPN
US-09-791-537-6910

  Query Match             14.1%;  Score 256.5;  DB 21;  Length 208;
  Best Local Similarity   33.8%;  Pred. No. 8.7e-14;
  Matches   77;  Conservative   29;  Mismatches   95;  Indels   27;  Gaps   13;

Qy      30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
           |  ||  :     || :|  || |  :| ||| | |      | |  | |:||| |||:
Db       1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 57

Qy      90 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 149
             | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db      58 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 110

Qy     150 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTL 207
           :|:  | | :  |||:       |:| :: :  ||:   :||   |||       || |
Db     111 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT----SNGSVF-- 160

Qy     208 DWTDNFDTFDGSRWGKGDWTFDGNRVDLT--DKNI-YSRDGMLILALT 252
             |      |: | | | |   :|   : |    |: :: || | | ||
Db     161 -WEPK-SYFNPSTWEKADGYSNGGVFNCTWRANNVNFTNDGKLKLGLT 206


RESULT 10
US-09-791-537-132244
```

```
; Sequence 132244, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132244
;   LENGTH: 394
;   TYPE: PRT
;   ORGANISM: pdb 1AXKA
US-09-791-537-132244

  Query Match             13.8%;  Score 252.5;  DB 21;  Length 394;
  Best Local Similarity   25.9%;  Pred. No. 4.9e-13;
  Matches   88;  Conservative   38;  Mismatches  133;  Indels   81;  Gaps   13;

Qy      30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
           |  ||  :     || :|  || |  :| ||| | |      | |  | |:||| |||:
Db       1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 57

Qy      90 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 149
             | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db      58 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 110

Qy     150 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKV---------YKYTPGQ 198
           :|:  | | :  |||:       |:| :: :  ||:   :|||            :| |
Db     111 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWVKYTSNASTDYWQNWTDGG 169

Qy     199 G------EGGSDFTLDWTDNFDTFDGSRWGKGD-----------WTFDGN---------- 231
           |        | :::::|::  :   |  |  |             |  :||
Db     170 GIVNAVNGSGGNYSVNWSNTGNFVVGKGWTTGSPFRTINYNAGVWAPNGNGYLTLYGWTR 229

Qy     232 --------------RVDLTDKNIYSRDGML--ILALTRKGQESFNGQVPRDDEPAPQSS 274
                        |   | |     ||    |  ||    | :|      :
Db     230 SPLIEYYVVDSWGTYRPTGTYKGTVKSDGGTYDIYTTTRYNAPSIDGDRTTFTQYWSVRQ 289

Qy     275 SSAPASSSSVPASSSSVPASSSSAFVPPSSSSATNAIHGM 314
           |  |  |::    |: | |  |               |||
Db     290 SKRPTGSNATITFSNHVNAWKS--------------HGM 314


RESULT 11
US-09-791-537-27063
; Sequence 27063, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27063
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-791-537-27063

  Query Match             13.7%;  Score 250.5;  DB 21;  Length 242;
  Best Local Similarity   35.5%;  Pred. No. 3.6e-13;
  Matches   61;  Conservative   25;  Mismatches   75;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |  |  :::   || :| ||| |  :| ||| | |   :   || || |:||
```

```
Db      80 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDI 136

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :
Db     137 EFLGKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 192

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           : |   |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     193 KHTATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 12
US-09-791-537-103347
; Sequence 103347, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103347
;   LENGTH: 239
;   TYPE: PRT
;   ORGANISM: synthetic construct
US-09-791-537-103347

  Query Match             13.7%;  Score 249;  DB 21;  Length 239;
  Best Local Similarity   36.0%;  Pred. No. 4.9e-13;
  Matches   63;  Conservative   26;  Mismatches   70;  Indels   16;  Gaps    8;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db      77 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 133

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||  :
Db     134 EFLGKDTTKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPGYIKWYVDG-VL 189

Qy     143 RKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           : |     :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     190 KHT---ATNIPTTPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 13
US-09-791-537-103324
; Sequence 103324, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 103324
;   LENGTH: 239
;   TYPE: PRT
;   ORGANISM: synthetic construct
US-09-791-537-103324

  Query Match             13.6%;  Score 248;  DB 21;  Length 239;
  Best Local Similarity   35.4%;  Pred. No. 6e-13;
  Matches   62;  Conservative   27;  Mismatches   70;  Indels   16;  Gaps    8;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db      77 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 133

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  ||     ||   :   | : ||||  :| | |::| |||  :
Db     134 EFLGKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VL 189

Qy     143 RKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           : |     :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     190 KHT---ATNIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 237


RESULT 14
US-09-791-537-73450
; Sequence 73450, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73450
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-791-537-73450

  Query Match             13.6%;  Score 248;  DB 21;  Length 242;
  Best Local Similarity   36.0%;  Pred. No. 6.1e-13;
  Matches   63;  Conservative   25;  Mismatches   75;  Indels   12;  Gaps    7;

Qy      20 ALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVE 79
           |||:    | |   |  :::   || :| ||| |  :| ||| | |   :   || || |
Db      78 ALTSPAYNK-FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDE 133

Qy      80 VDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 139
           :||| |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||
Db     134 IDIEFLGKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDG 190

Qy     140 QEVRKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           | :: |   |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     191 Q-LKHTATNQIPTTLGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 15
US-09-791-537-79889
; Sequence 79889, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 79889
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-791-537-79889

  Query Match             13.6%;  Score 248;  DB 21;  Length 242;
  Best Local Similarity   36.0%;  Pred. No. 6.1e-13;
  Matches   63;  Conservative   25;  Mismatches   75;  Indels   12;  Gaps    7;
```

```
Qy      20 ALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVE 79
           |||:    | |   |  :::   || :| ||| |  :| ||| | |   :   || || |
Db      78 ALTSPAYNK-FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDE 133

Qy      80 VDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 139
           :||| |||:    | |  |  ||     ||   :   |  |:||| :| || ::| |||
Db     134 IDIEFLGKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDG 190

Qy     140 QEVRKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           | :: |   |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     191 Q-LKHTATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241
```

Search completed: January 9, 2003, 12:13:29
Job time : 165.402 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:14:27 ; Search time 16.5613 Seconds
(without alignments)
2025.857 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKKTAVKSALAVAAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1824	100.0	349	2	A44507	licheninase (EC 3.
2	268.5	14.7	851	2	H84053	endo-beta-1,3-1,4
3	265.5	14.6	252	2	A48378	licheninase (EC 3.
4	248	13.6	242	1	LXBS	licheninase (EC 3.
5	247.5	13.6	276	2	I40453	licheninase (EC 3.
6	242.5	13.3	239	1	A29091	licheninase (EC 3.
7	242.5	13.3	802	2	A36910	xylanase, beta(1,3
8	237	13.0	334	1	S23498	licheninase (EC 3.
9	233.5	12.8	243	1	S15388	licheninase (EC 3.
10	230	12.6	507	2	S64507	probable membrane
11	228	12.5	237	1	S11927	licheninase (EC 3.
12	227.5	12.5	238	1	S19012	licheninase (EC 3.
13	224.5	12.3	242	2	JS0611	endo-1,3(4)-beta-g
14	221	12.1	246	2	D97245	endo-1,3(4)-beta-g
15	206	11.3	263	2	AB3055	endo-1,3-1,4-beta-
16	206	11.3	289	2	C98231	endo-1,3-1,4-beta-
17	197	10.8	269	2	H95976	endo-beta-1,3-1,4-
18	190.5	10.4	302	2	G84053	hybrid-endo-beta-1
19	189.5	10.4	467	2	S30839	UTR2 protein - yea
20	180.5	9.9	268	2	S34804	exoK protein - Rhi
21	173	9.5	282	2	T02354	xyloglucan endo-1,
22	173	9.5	642	2	B72428	laminarinase - The
23	172.5	9.5	277	2	B85354	hypothetical prote
24	170.5	9.3	282	2	A85354	hypothetical prote
25	164	9.0	286	2	S71225	xyloglucan endo-1,
26	159.5	8.7	287	2	T04236	xyloglucan endo-1,
27	157.5	8.6	269	2	S61555	xyloglucan endo-1,
28	154.5	8.5	277	2	S71222	xyloglucan endo-1,
29	149	8.2	289	2	T06166	xyloglucan endotra
30	146.5	8.0	284	2	T52097	xyloglucan endo-1,
31	145.5	8.0	310	2	A86239	protein T10024.17
32	145	7.9	305	2	G84568	probable xylogluca
33	140.5	7.7	280	2	T02090	xyloglucan endo-1,
34	139.5	7.6	286	2	T06202	xyloglucan endo-1,
35	138	7.6	422	2	S48564	probable membrane
36	137	7.5	292	2	T06201	xyloglucan endo-1,
37	133.5	7.3	286	2	S48201	licheninase (EC 3.
38	132	7.2	295	2	T10210	xyloglucan endo-1,
39	131.5	7.2	301	2	C87296	beta-glucanase [im
40	129.5	7.1	1324	2	T18265	endo-1,3(4)-beta-g
41	128.5	7.0	277	2	JE0156	end-xyloglucan tra
42	128.5	7.0	292	2	T04514	xyloglucan endo-1,
43	128	7.0	283	2	T07678	xyloglucan endo-1,
44	128	7.0	846	2	AD2672	conserved hypothet
45	128	7.0	848	2	B97454	hypothetical prote

ALIGNMENTS

RESULT 1
A44507
licheninase (EC 3.2.1.73) - Fibrobacter succinogenes
C;Species: Fibrobacter succinogenes
C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 08-Oct-1999
C;Accession: A44507
R;Teather, R.M.; Erfle, J.D.
J. Bacteriol. 172, 3837-3841, 1990
A;Title: DNA sequence of a Fibrobacter succinogenes mixed-linkage beta-glucanase (1,3
A;Reference number: A44507; MUID:90299807; PMID:2193918
A;Accession: A44507
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TEA>
A;Cross-references: EMBL:M33676; NID:g148575; PIDN:AAA24896.1; PID:g148576
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 100.0%; Score 1824; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.5e-122;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60

Qy      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120

Qy     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180

Qy     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 240

Qy     241 YSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 YSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV 300

Qy     301 PPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNGHKRYRVNFEH 349
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 PPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNGHKRYRVNFEH 349
```

RESULT 2
H84053
endo-beta-1,3-1,4 glucanase (licheninase) bglS [imported] - Bacillus halodurans (stra
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: H84053

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H84053
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-851 <STO>
A;Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB06951.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: bglS

Query Match 14.7%; Score 268.5; DB 2; Length 851;
Best Local Similarity 29.4%; Pred. No. 2.6e-11;
Matches 83; Conservative 32; Mismatches 112; Indels 55; Gaps 12;

```
Qy      21 LTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEV 80
           |  :  |  :   || | :  ||| ||  || |   ||||| | |  |    || || |:
Db      93 LENDEQAPPYKAGELRTNDFYQYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEI 151

Qy      81 DIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQ 140
           ||| |||:    | |  |    |    :| :: :   | ::|:||  ||  : : | |:|:
Db     152 DIEFLGKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGE 208

Qy     141 EVRKTEGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP-- 196
             |         |:  | | :  |||       | | ||    ||     ||   :|||
Db     209 AVHTA----TENIPQTPQKIMMNLWPGVGVDEWTGVFDGDNTPLHADYEWV---RYTPLE 261

Qy     197 --GQGEGGSDFTLDWT----------------------DNFDTFDGSRWG-KGDWTFDG 230
             : | | |   :                        : |:||:   |     || :|
Db     262 ALDEEESGDDNEEEPVEEVEEEEPAEDEEVSVRPIGSAIYETFNTFNEDIWSIAHGWT-NG 320

Qy     231 NRVDLT--DKNIYSRDGMLILALTRKGQESFNGQVPRDDEPA 270
           : |  |  |   :|::  |: ::          ||| |
Db     321 QMFNATWYDSQITFSNGIMRFAIDKE----------DDEEA 351
```

RESULT 3
A48378
licheninase (EC 3.2.1.73) - Bacillus brevis
N;Alternate names: lichenase
C;Species: Bacillus brevis
C;Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 07-May-1999
C;Accession: A48378
R;Louw, M.E.; Reid, S.J.; Watson, T.G.
Appl. Microbiol. Biotechnol. 38, 507-513, 1993
A;Title: Characterization, cloning and sequencing of a thermostable endo-(1,3-1,4) beta-
A;Reference number: A48378; MUID:93159752; PMID:7763386
A;Accession: A48378
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-252 <LOU>
A;Note: sequence extracted from NCBI backbone (NCBIN:124672, NCBIP:124673)
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 14.6%; Score 265.5; DB 2; Length 252;
Best Local Similarity 35.8%; Pred. No. 9.2e-12;
Matches 62; Conservative 24; Mismatches 74; Indels 13; Gaps 6;

```
Qy      26 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 85
           ||:::   || | :  || ||  || |   ||||| | | . |    || || |:||| |
Db      90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy      86 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 145
           ||:    | |  |    :| :: :   | ::|:||  ||  : : | |:|: |
Db     149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy     146 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 196
             |:  | | :  |||       | | ||    |:: : :||   :|||
Db     206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTP 251
```

RESULT 4
LXBS
licheninase (EC 3.2.1.73) precursor - Bacillus subtilis
N;Alternate names: 1,3-1,4-beta-D-glucan 4-gluconohydrolase; beta-glucanase; lichenas
N;Contains: licheninase E-1; licheninase E-2
C;Species: Bacillus subtilis
C;Date: 30-Jun-1988 #sequence_revision 13-Mar-1998 #text_change 16-Jun-2000
C;Accession: I40370; B69594; A22914; A90026; A90027; JU0110
R;Wolf, M.; Geczi, A.; Simon, O.; Borriss, R.
Microbiology 141, 281-290, 1995
A;Title: Genes encoding xylan and beta-glucan hydrolysing enzymes in Bacillus subtili
A;Reference number: I40370; MUID:95219081; PMID:7704256
A;Accession: I40370
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-242 <RES>
A;Cross-references: EMBL:Z46862; NID:g599673; PIDN:CAA86922.1; PID:g599674
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: B69594
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-242 <KUN>
A;Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB15943.1; PID:g26364
A;Experimental source: strain 168
R;Murphy, N.; McConnell, D.J.; Cantwell, B.A.
Nucleic Acids Res. 12, 5355-5367, 1984
A;Title: The DNA sequence of the gene and genetic control sites for the excreted B. s
A;Reference number: A93526; MUID:84272222; PMID:6087283
A;Accession: A22914
A;Molecule type: DNA
A;Residues: 1-203,'L',205-242 <MUR>
A;Cross-references: EMBL:X00754; NID:g39818; PIDN:CAA25328.1; PID:g685236
A;Experimental source: strain C120
A;Note: the authors translated the codon CAA for residue 29 as Lys and CCA for residu
R;Tezuka, H.; Yuuki, T.; Yabuuchi, S.
Agric. Biol. Chem. 53, 2335-2339, 1989
A;Title: Construction of a beta-glucanase hyperproducing Bacillus subtilis using the
A;Reference number: A90026
A;Accession: A90026
A;Molecule type: DNA
A;Residues: 1-23,'S',25-82,'S',84-242 <TEZ>
A;Cross-references: DDBJ:D00518; NID:g216243; PIDN:BAA00405.1; PID:g216244
A;Experimental source: strain Y-25, clone pLE100
R;Yuuki, T.; Tezuka, H.; Yabuuchi, S.
Agric. Biol. Chem. 53, 2341-2346, 1989
A;Title: Purification and some properties of two enzymes from a beta-glucanase hyperp
A;Reference number: A90027
A;Contents: annotation
A;Note: source was hyperproducing strain HL-25 with gene from strain Y-25
A;Note: the amino ends of the mature forms of E-1 and E-2 are pyroglutamic acid and g
C;Genetics:
A;Gene: bglS
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin an
C;Superfamily: licheninase
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F;1-28/Domain: signal sequence #status predicted <SIG>

```
F;29-242/Product: licheninase #status predicted <MAT>
F;29/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) (partial) #status

  Query Match             13.6%;  Score 248;  DB 1;  Length 242;
  Best Local Similarity   36.0%;  Pred. No. 1.5e-10;
  Matches   63;  Conservative   25;  Mismatches   75;  Indels   12;  Gaps    7;

Qy      20 ALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVE 79
           |||:    | |   |  :::   || :| ||| |  :| ||| | |   :   || || |
Db      78 ALTSPAYNK-FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDE 133

Qy      80 VDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 139
           :||| |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||
Db     134 IDIEFLGKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDG 190

Qy     140 QEVRKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           | :: |   |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     191 Q-LKHTATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 5
I40453
licheninase (EC 3.2.1.73) - Bacillus sp.
C;Species: Bacillus sp.
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 22-Jun-1999
C;Accession: I40453; S32688
R;Tabernero, C.; Coll, P.M.; Fernandez-Abalos, J.M.; Perez, P.; Santamaria, R.I.
Appl. Environ. Microbiol. 60, 1213-1220, 1994
A;Title: Cloning and DNA sequencing of bgaA, a gene encoding an endo-beta-1,3-1,4-glucan
A;Reference number: I40453; MUID:94288605; PMID:7517127
A;Accession: I40453
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-276 <RES>
A;Cross-references: EMBL:Z12151; NID:g296931; PIDN:CAA78135.1; PID:g296932
C;Genetics:
A;Gene: bgal
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase

  Query Match             13.6%;  Score 247.5;  DB 2;  Length 276;
  Best Local Similarity   36.1%;  Pred. No. 2e-10;
  Matches   61;  Conservative   20;  Mismatches   75;  Indels   13;  Gaps    6;

Qy      30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
           :   || | :   || ||  || | ::||||| | |  |    :  || |:||| |||:
Db      96 YKAGELRTNDFYHYGLFEVSMKPAKSTGTVSSFFTY-TGPWDWENDPWDEIDIEFLGKDT 154

Qy      90 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 149
             | |  |   |    :| :| :   |   |:||  || |  :|| |:|: |
Db     155 TKIQFNYFTNGVGG---NEHYHELGFDAADDFNTYAFEWRPESIRWFVNGELVHTA---- 207

Qy     150 VSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 196
           |:  | | :  |||       | |:|:    |:    :||   ||||
Db     208 TENIPQTPQKIMMNLWPGIGVDGWTGRFNGEDTPVVTQYDWV---KYTP 253


RESULT 6
A29091
licheninase (EC 3.2.1.73) beta - Bacillus amyloliquefaciens
N;Alternate names: 1,3-1,4-beta-D-glucan 4-glucanohydrolase; beta-glucanase; lichenase
C;Species: Bacillus amyloliquefaciens
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29091
R;Hofemeister, J.; Kurtz, A.; Borriss, R.; Knowles, J.
Gene 49, 177-187, 1986
A;Title: The beta-glucanase gene from Bacillus amyloliquefaciens shows extensive homolog
A;Reference number: A91564; MUID:87192007; PMID:3106158
A;Accession: A29091
A;Molecule type: DNA
A;Residues: 1-239 <HOF>
A;Cross-references: GB:M15674; NID:g143009; PIDN:AAA87323.1; PID:g143010
A;Experimental source: strain BE20/78
C;Genetics:
A;Gene: bglA
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match             13.3%;  Score 242.5;  DB 1;  Length 239;
  Best Local Similarity   34.3%;  Pred. No. 3.7e-10;
  Matches   59;  Conservative   27;  Mismatches   75;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db      77 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 133

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :
Db     134 EFLGKDTTKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 189

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           : |   |:    |   :  |||:      |:| ::    |::   :|::  |
Db     190 KHTATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYDWMRYRK 238


RESULT 7
A36910
xylanase, beta(1,3-1,4)-glucanase - Ruminococcus flavefaciens
C;Species: Ruminococcus flavefaciens
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
C;Accession: A36910
R;Flint, H.J.; Martin, J.; McPherson, C.A.; Daniel, A.S.; Zhang, J.X.
J. Bacteriol. 175, 2943-2951, 1993
A;Title: A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-glucanase do
A;Reference number: A36910; MUID:93259938; PMID:8491715
A;Accession: A36910
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-802 <FLI>
A;Cross-references: GB:S61204; NID:g385910; PIDN:AAB26620.1; PID:g385911
A;Note: sequence extracted from NCBI backbone (NCBIN:131871, NCBIP:131872)
C;Superfamily: endo-1,4-beta-xylanase homology; Thermotoga xylanase A amino-terminal
F;42-239/Domain: endo-1,4-beta-xylanase homology <XYL>
F;259-401/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

  Query Match             13.3%;  Score 242.5;  DB 2;  Length 802;
  Best Local Similarity   34.9%;  Pred. No. 1.7e-09;
  Matches   58;  Conservative   21;  Mismatches   74;  Indels   13;  Gaps    6;

Qy      30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
           :|| |  |     || :|  |:     | ||| | |   |   |  || |:|||:||||
Db     634 YSGGEFRTNNFYHYGYYECSMQAMKNDGVVSSFFTYTGPS---DDNPWDEIDIEILGKNT 690

Qy      90 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 149
             |  |   |     || : :   : :|:|||| :| |||: | |||:|| :
Db     691 TQVQFNYYTNGQGKH---EKLYDLGFDSSEAYHTYGFDWQPNYIAWYVDGREVYRA---- 743

Qy     150 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
             ::  | | :  | |   :   |:  |:  : ||     ||   |
Db     744 TQDIPKTPGKIMMNAWPGLTVDDWLKAFN-GRTPLTAHYQWVTYNK 788


RESULT 8
S23498
licheninase (EC 3.2.1.73) licB precursor - Clostridium thermocellum
N;Alternate names: beta-1,3-1,4-glucanase licB; lichenase licB
C;Species: Clostridium thermocellum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S23498; S22137
R;Schimming, S.; Schwarz, W.H.; Staudenbauer, W.L.
Eur. J. Biochem. 204, 13-19, 1992
A;Title: Structure of the Clostridium thermocellum gene licB and the encoded beta-1,3
1 cellulases.
A;Reference number: S23498; MUID:92155194; PMID:1740123
A;Accession: S23498
A;Molecule type: DNA
```

```
A;Residues: 1-334 <SCH>
A;Cross-references: EMBL:X63355; NID:g40697; PIDN:CAA44959.1; PID:g40698
C;Genetics:
A;Gene: licB
C;Superfamily: licheninase licB; Clostridium cellulase repeat homology
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-334/Product: licheninase #status predicted <MAT>
F;273-296/Domain: Clostridium cellulase repeat homology <CCR1>
F;308-331/Domain: Clostridium cellulase repeat homology <CCR2>

  Query Match             13.0%;  Score 237;  DB 1;  Length 334;
  Best Local Similarity   30.9%;  Pred. No. 1.4e-09;
  Matches   80;  Conservative   27;  Mismatches   94;  Indels   58;  Gaps   13;

Qy      28 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 81
           ::: |:  |   | :       || :| ||| |   | ||| | |   |   |  || |:|
Db      82 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 138

Qy      82 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 141
           || |||:    | |      |    :|  | :   | | ||||| || |:|: : |||::
Db     139 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 195

Qy     142 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVY------- 192
           | :       |: | | :  |||       |:|::|  : ||     :|| |
Db     196 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPLQAEYEYVKYYPNGVPQD 250

Qy     193 ---------KYTP-----------GQGE-GGSDFT------LDWTDNFDTFDGSRWGKG 224
                    ||          | |    ||::      |   | |   | |
Db     251 NPTPTPTIAPSTPTNPNLPLKGDVNGDGHVNSSDYSLFKRYLLRVIDRFPVGDQS---VA 307

Qy     225 DWTFDGNRVDLTDKNIYSR 243
           |   || |:| ||  :  |
Db     308 DVNRDG-RIDSTDLTMLKR 325


RESULT 9
S15388
licheninase (EC 3.2.1.73) - Bacillus licheniformis
N;Alternate names: endo-beta-1,3-1,4-D-glucanase; lichenase
C;Species: Bacillus licheniformis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S15388
R;Lloberas, J.; Perez-Pons, J.A.; Querol, E.
Eur. J. Biochem. 197, 337-343, 1991
A;Title: Molecular cloning, expression and nucleotide sequence of the endo-beta-1,3-1,4-
A;Reference number: S15388; MUID:91224124; PMID:2026156
A;Accession: S15388
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-243 <LLO>
A;Cross-references: EMBL:X57279; NID:g39558; PIDN:CAA40547.1; PID:g39559
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match             12.8%;  Score 233.5;  DB 1;  Length 243;
  Best Local Similarity   33.7%;  Pred. No. 1.7e-09;
  Matches   58;  Conservative   26;  Mismatches   77;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :|  || |   | ||| | |   :   || || |:||
Db      81 TSPSYNKFDCGENRSVQTYGYGLYEVNMKPAKNVGIVSSFFTYTGPT---DGTPWDEIDI 137

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |   |     ||   :   |  ::||| :| || ::| |||| :
Db     138 EFLGKDTTKVQFNYYTNGVG---NHEKIVNLGFDAANSYHTYAFDWQPNSIKWYVDGQ-L 193

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           : |   |:    |   :  |||:      |:| ::    || : ::||:  |
Db     194 KHTATTQIPQTPGK--IMMNLWNGAGVDEWLGSYN-GVTPLSRSLHWVRYTK 242


RESULT 10
```

```
S64507
probable membrane protein YGR189c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G7553
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
C;Accession: S64507
R;Arroyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M.I.; Sanchez-Perez, M.; Nombela, C.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64499
A;Accession: S64507
A;Molecule type: DNA
A;Residues: 1-507 <ARR>
A;Cross-references: EMBL:Z72974; NID:g1323335; PID:e243566; PID:g1323336; GSPDB:GN000
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:CRH1; MIPS:YGR189c
A;Cross-references: SGD:S0003421
A;Map position: 7R
C;Keywords: transmembrane protein
F;6-22/Domain: transmembrane #status predicted <TMM>

  Query Match             12.6%;  Score 230;  DB 2;  Length 507;
  Best Local Similarity   26.9%;  Pred. No. 7.5e-09;
  Matches   87;  Conservative   57;  Mismatches  134;  Indels   46;  Gaps   13;

Qy       6 TAVKSALAVAAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLY 65
           | :| |  :   :  |:  : || :    | :   : ||| |  :| |  :| ||| :|
Db      70 TDLKHAGEIKYGSDGLSMTL-AKRYDNPSLKSNFYIMYGKLEVILKAANGTGIVSSFYLQ 128

Qy      66 QNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYG 125
             :  :        |:||| :| :   ||||  :          : | |    |: || |
Db     129 SDDLD--------EIDIEWVGGDNTQFQSNFFSKGDTTTYDRGEFHGVDTPTDK-FHNYT 179

Qy     126 LEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQG-------LRFNLWSS---ESAA----W 171
           |:|  :   | :||: ||      | : | ::|       |   :|:    ::||    |
Db     180 LDWAMDKTTWYLDGESVR------VLSNTSSEGYPQSPMYLMMGIWAGGDPDNAAGTIEW 233

Qy     172 V-GQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDG 230
             |: : :  |   :|  | |  |: |:     | :  | :: :   || :|: |   :
Db     234 AGGETNYNDAPFTMYIEKVIVTDYSTGKKYTYGDQSGSW-ESIEADGGSIYGRYDQAQE- 291

Qy     231 NRVDLTDKNIYSRDGMLILALTRKG--QESFNGQVPRDDEPAPQSSSSAPASS--SSVPA 286
                |  : :  | :  : |       | : |          ||:|:  ||  ||  :
Db     292 ------DFAVLANGGSISSSSTSSSTVSSSASSTVSSSVSSTVSSSASSTVSSSVSSTVS 345

Qy     287 SSSSVPASSSSAFVPPSSSSATNA 310
           |||||  :|||::   ||||:||::
Db     346 SSSSVSSSSSTS---PSSSTATSS 366


RESULT 11
S11927
licheninase (EC 3.2.1.73) precursor [validated] - Bacillus macerans
N;Alternate names: endo-beta-1,3-1,4-glucanase; lichenase
C;Species: Bacillus macerans
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2000
C;Accession: S11927
R;Borriss, R.; Buettner, K.; Maentsaelae, P.
Mol. Gen. Genet. 222, 278-283, 1990
A;Title: Structure of the beta-1,3-1,4-glucanase gene of Bacillus macerans: homologie
A;Reference number: S11927; MUID:91109712; PMID:2274030
A;Accession: S11927
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <BOR>
A;Cross-references: EMBL:X55959; NID:g296715; PIDN:CAA39426.1; PID:g296716
R;Hahn, M.; Heinemann, U.
submitted to the Brookhaven Protein Data Bank, December 1994
A;Reference number: A67074; PDB:1MAC
A;Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 26-37,'P',39-2
A;Note: recombinant form with residues 26-41 from Bacillus amyloliquefaciens expresse
R;Keitel, T.; Simon, O.; Borriss, R.; Heinemann, U.
Proc. Natl. Acad. Sci. U.S.A. 90, 5287-5291, 1993
```

```
A;Title: Molecular and active-site structure of a Bacillus 1,3-1,4-beta-glucanase.
A;Reference number: A47562; MUID:93281743; PMID:8099449
A;Contents: annotation; X-ray crystallography, 2.3 angstroms
A;Note: recombinant form with residues 26-41 from Bacillus amyloliquefaciens expressed i
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin and c
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-25/Domain: signal sequence #status predicted <PRO>
F;26-237/Product: licheninase #status experimental <MAT>
F;55-84/Disulfide bonds: #status experimental
F;128/Active site: Glu #status predicted

  Query Match             12.5%;  Score 228;  DB 1;  Length 237;
  Best Local Similarity   34.3%;  Pred. No. 4e-09;
  Matches   60;  Conservative   25;  Mismatches   74;  Indels   16;  Gaps    8;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: :   |  ||  :     || :|  || |  :| ||| | |      | |  | |:||
Db      75 TSSAYNKFDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDI 131

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |    |     ||  ::   | : |||| :| | |::| |||  :
Db     132 EFLGKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VL 187

Qy     143 RKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           : |     :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     188 KHT---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 235


RESULT 12
S19012
licheninase (EC 3.2.1.73) precursor - Bacillus polymyxa
N;Alternate names: endo-beta-1,3-1,4-glucanase; lichenase
C;Species: Bacillus polymyxa
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 18-Jun-1999
C;Accession: S19012
R;Gosalbes, M.J.; Perez-Gonzalez, J.A.; Gonzalez, R.; Navarro, A.
J. Bacteriol. 173, 7705-7710, 1991
A;Title: Two beta-glycanase genes are clustered in Bacillus polymyxa: molecular cloning,
A;Reference number: S19011; MUID:92041687; PMID:1938968
A;Accession: S19012
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-238 <GOS>
A;Cross-references: EMBL:X57094; NID:g48815; PIDN:CAA40379.1; PID:g48817
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin and c
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-26/Domain: signal sequence #status predicted <PRO>
F;27-237/Product: licheninase #status predicted <MAT>
F;56-85/Disulfide bonds: #status predicted
F;129/Active site: Glu #status predicted

  Query Match             12.5%;  Score 227.5;  DB 1;  Length 238;
  Best Local Similarity   34.3%;  Pred. No. 4.4e-09;
  Matches   61;  Conservative   26;  Mismatches   74;  Indels   17;  Gaps    9;

Qy      20 ALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVE 79
           :||:  : | |   |  :     || :|  || |  :| ||| | |   |    |  | | |
Db      74 SLTSPANNK-FDCGEYRSTNNYGYGLYEVSMKPAKNTGIVSSFFTYTGPSH---GTQWDE 129

Qy      80 VDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 139
           :||| |||:    | |  |    |     ||   :   |  :|||| :| | |::| |||
Db     130 IDIEFLGKDTTKVQFNYYTNGVGGH---EKIINLGFDASTSFHTYAFDWQPGYIKWYVDG 186

Qy     140 QEVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           :: |     :|:  | | :  |||:      :|:| :: :  ||:   :||   |||
Db     187 -VLKHT---ATTNIPSTPGKIMMNLWNGTGVDSWLGSYNGAN-PLYAEYDWV---KYT 236


RESULT 13
JS0611
endo-1,3(4)-beta-glucanase (EC 3.2.1.6) precursor - Clostridium thermocellum
N;Alternate names: laminarinase
C;Species: Clostridium thermocellum
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jan-1997
C;Accession: JS0611; S18726
R;Zverlov, V.V.; Laptev, D.A.; Tishkov, V.I.; Velikodvorskaja, G.A.
Biochem. Biophys. Res. Commun. 181, 507-512, 1991
A;Title: Nucleotide sequence of the Clostridium thermocellum laminarinase gene.
A;Reference number: JS0611; MUID:92095946; PMID:1755832
A;Accession: JS0611
A;Molecule type: DNA
A;Residues: 1-242 <ZVE1>
A;Cross-references: EMBL:X58392
R;Zverlov, V.V.; Velikodvorskaya, G.A.
Biotechnol. Lett. 12, 811-816, 1990
A;Title: Cloning the Clostridium thermocellum thermostable laminarinase gene in Esche
A;Reference number: S18726
A;Accession: S18726
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-242 <ZVE2>
A;Cross-references: EMBL:X58392
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991
C;Genetics:
A;Gene: lam1
C;Superfamily: licheninase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-242/Product: endo-1,3(4)-beta-glucanase #status predicted <END>

  Query Match             12.3%;  Score 224.5;  DB 2;  Length 242;
  Best Local Similarity   33.5%;  Pred. No. 7.3e-09;
  Matches   62;  Conservative   25;  Mismatches   71;  Indels   27;  Gaps    8;

Qy      14 VAAAAAALTTNVS--------AKDFSGAELYTLEEVQ------YGKFEARMKMAAASGTV 59
           |:    | | ::|         ::: |:  |   | :      || :| ||| |   | |
Db      58 VSTVLEAFTGDISNGKMILTLDREYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIV 117

Qy      60 SSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQ 119
           || | |   |   |  || |:||| |||:    | |      |    :|  | :  | |
Db     118 SSFFTYTGPS---DNNPWDEIDIEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQ 171

Qy     120 AFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDE 177
           |||||  || |:|: : |||::| :       |:  | | :  |||       |:|::|
Db     172 DFHTYGFEWRPDYIDFYVDGKKVYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD- 226

Qy     178 SKLPL 182
           : ||
Db     227 GRTPL 231


RESULT 14
D97245
endo-1,3(4)-beta-glucanase family 16 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: D97245
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97245
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-246 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80751.1; PID:g15025848; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2807
C;Superfamily: licheninase
```

```
  Query Match            12.1%;  Score 221;  DB 2;  Length 246;
  Best Local Similarity  33.7%;  Pred. No. 1.3e-08;
  Matches   58;  Conservative   19;  Mismatches   73;  Indels   22;  Gaps    8;

Qy      30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
           ::| |  :     || :   || |   |  || | |   |   |  || |:||| |||:
Db      91 YAGGEYRSNNRYGYGLYRVSMKPAKHIGVDSSFFSYTGPS---DNNPWDEIDIEFLGKDT 147

Qy      90 GSFQSNIITGKAGAQKTSEKH---HAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 146
             | |  |   |      ||   : :   | : ||||  |  ||: | |||:|| :
Db     148 TEVQFNYYTNGVG------KHEYLYKLGFDASKGFHTYGYIWEQNYIAWLVDGKEVYRA- 200

Qy     147 GGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 196
           ||:    | :  |||      :|:| :|    |:  : ||     | |
Db     201 ---TSNIPTHPGKVMMNLWPGIGVDSWLGAYD-GVTPVKAYYNWA---MYNP 245


RESULT 15
AB3055
endo-1,3-1,4-beta-glycanase exoK [imported] - Agrobacterium tumefaciens (strain C58, Dup
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C;Accession: AB3055
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB3055
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44856.1; PID:g17742502; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: exoK
A;Map position: linear chromosome
C;Superfamily: licheninase

  Query Match            11.3%;  Score 206;  DB 2;  Length 263;
  Best Local Similarity  35.1%;  Pred. No. 1.7e-07;
  Matches   54;  Conservative   26;  Mismatches   56;  Indels   18;  Gaps    6;

Qy      28 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 87
           ::|:  |: |    :|| :|||||  |  ||  |: | |   :   | :|  |:| |||||
Db      89 RNFACGEIQTKGRYRYGTYEARMKAATGSGLNSAFFTYIGPT---DKKPHDEIDFEVLGK 145

Qy      88 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 146
           | |  | |  |  | |    :||   |   ||  |: |   | |  :|: |:|: |
Db     146 NTGKVQLNQYIAAKGG----NEKLVPVEGGADAGFNDYAFVWEPQRLRYYVNGKLVH--- 198

Qy     147 GGQVSNLT----GTQGLRFNLWSSESAA-WVGQF 175
           :|:: |      | : |:|| :::   |:|:|
Db     199 --EVTDETKIPQNAQKIFFSLWGTDTLKDWMGKF 230


Search completed: January  9, 2003, 12:19:01
Job time : 18.5613 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:06:42 ; Search time 8.88657 Seconds
(without alignments)
1628.889 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKKTAVKSALAVAAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1824	100.0	349	1	GUB_FIBSU	P17989 fibrobacter
2	266.5	14.6	259	1	GUB_BACBR	P37073 bacillus br
3	248	13.6	242	1	GUB_BACSU	P04957 bacillus su
4	242.5	13.3	239	1	GUB_BACAM	P07980 bacillus am
5	242.5	13.3	802	1	XYND_RUMFL	Q53317 ruminococcu
6	237	13.0	334	1	GUB_CLOTM	P29716 clostridium
7	233.5	12.8	243	1	GUB_BACLI	P27051 bacillus li
8	230	12.6	507	1	YG46_YEAST	P53301 saccharomyc
9	228	12.5	237	1	GUB_PAEMA	P23904 paenibacill
10	227.5	12.5	238	1	GUB_PAEPO	P45797 paenibacill
11	197	10.8	269	1	EXOK_RHIME	P33693 rhizobium m
12	189.5	10.4	347	1	UTR2_YEAST	P32623 saccharomyc
13	157.5	8.6	269	1	MER5_ARATH	P24806 arabidopsis
14	133.5	7.3	286	1	GUB_RHOMR	P45798 rhodothermu
15	128	7.0	283	1	BRU1_SOYBN	P35694 glycine max
16	119	6.5	465	1	EXSH_RHIME	O33680 rhizobium m
17	117	6.4	397	1	CGKA_ALTCA	P43478 alteromonas
18	115	6.3	682	1	E13B_BACCI	P23903 bacillus ci
19	108.5	5.9	989	1	SERA_PLAFG	P13823 plasmodium
20	107.5	5.9	719	1	PBPA_STRPN	Q04707 streptococc
21	107	5.9	484	1	YOD0_YEAST	Q08193 saccharomyc
22	104	5.7	962	1	GUNA_PSEFL	P10476 pseudomonas
23	103.5	5.7	1580	1	ACC8_HUMAN	Q09428 homo sapien
24	102.5	5.6	879	1	GUNI_CLOTM	Q02934 clostridium
25	102	5.6	465	1	EGLC_RHIME	Q9z3q2 rhizobium m
26	102	5.6	650	1	SCO2_NEUCR	Q01277 neurospora
27	101	5.5	691	1	VIV1_MAIZE	P26307 zea mays (m
28	100.5	5.5	1036	1	NIT2_NEUCR	P19212 neurospora
29	100	5.5	748	1	GUNC_PSEFL	P27033 pseudomonas
30	99	5.4	344	1	GUN4_TRIRE	O14405 trichoderma
31	99	5.4	525	1	Y487_TREPA	O83500 treponema p
32	99	5.4	560	1	YDEN_ECOLI	P77318 escherichia
33	98.5	5.4	658	1	GUN3_FIBSU	P14250 fibrobacter
34	98.5	5.4	1655	1	OMPB_RICCN	Q9kka3 r outer mem
35	98	5.4	1845	1	Z236_HUMAN	Q9ul36 homo sapien
36	97.5	5.3	1385	1	C5AA_BACUD	Q45760 bacillus th
37	97.5	5.3	1581	1	ACC8_CRICR	Q09427 cricetus cr
38	97.5	5.3	1581	1	ACC8_RAT	Q09429 rattus norv
39	97	5.3	1770	1	PMPC_CHLTR	O84419 chlamydia t
40	96.5	5.3	526	1	CAP_YEAST	P17555 saccharomyc
41	96	5.3	371	1	LIG2_PHACH	P49012 phanerochae
42	95.5	5.2	872	1	IP3L_RAT	P42335 rattus norv
43	95.5	5.2	1289	1	C5AB_BACUD	Q45753 bacillus th
44	95	5.2	400	1	GUN5_BACAG	O85465 bacillus ag
45	94.5	5.2	451	1	MYBH_DICDI	P34127 dictyosteli

ALIGNMENTS

RESULT 1
GUB_FIBSU
ID GUB_FIBSU STANDARD; PRT; 349 AA.
AC P17989;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Mixed linkage beta-
DE glucanase) (Lichenase).
OS Fibrobacter succinogenes (Bacteroides succinogenes).
OC Bacteria; Fibrobacter/Acidobacteria group; Fibrobacter group;
OC Fibrobacter.
OX NCBI_TaxID=833;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-57.
RC STRAIN=Isolate S85;
RX MEDLINE=90299807; PubMed=2193918;
RA Teather R.M., Erfle J.D.;
RT "DNA sequence of a Fibrobacter succinogenes mixed-linkage
RT beta-glucanase (1,3-1,4-beta-D-glucan 4-glucanohydrolase) gene.";
RL J. Bacteriol. 172:3837-3841(1990).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC ----------

CC ----------
DR EMBL; M33676; AAA24896.1; -.
DR PIR; A44507; A44507.
DR HSSP; P23904; 1AJK.
DR InterPro; IPR000757; Glyco_hydro_16.
DR Pfam; PF00722; Glyco_hydro_16; 1.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Hydrolase; Glycosidase; Signal; Repeat.
FT SIGNAL 1 27
FT CHAIN 28 349 BETA-GLUCANASE.
FT ACT_SITE 79 79 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 83 83 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 271 307 5 X 7 AA TANDEM REPEATS OF P-X-S-S-S-S-X.
FT REPEAT 271 277 1.
FT REPEAT 278 284 2.
FT REPEAT 285 291 3.
FT REPEAT 292 298 4.
FT REPEAT 301 307 5.
SQ SEQUENCE 349 AA; 37737 MW; 16DC4F5BDEFC578A CRC64;

Query Match 100.0%; Score 1824; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.9e-123;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MNIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVS 60

Qy      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA 120

Qy     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 FHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKL 180

Qy     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 PLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 240

Qy     241 YSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 YSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV 300

Qy     301 PPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNGHKRYRVNFEH 349
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 PPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNGHKRYRVNFEH 349
```

RESULT 2
GUB_BACBR
ID GUB_BACBR STANDARD; PRT; 259 AA.
AC P37073;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN BGLBB.
OS Bacillus brevis.
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.
OX NCBI_TaxID=1393;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93159752; PubMed=7763386;
RA Louw M.E., Reid S.J., Watson T.G.;
RT "Characterization, cloning and sequencing of a thermostable
RT endo-(1,3-1,4) beta-glucanase-encoding gene from an alkalophilic
RT Bacillus brevis.";
RL Appl. Microbiol. Biotechnol. 38:507-513(1993).
CC -!- FUNCTION: HYDROLYSES B-GLUCANS CONTAINING MIXED BETA-1,3 AND
CC BETA-1,4 LINKAGES. THIS ENZYME IS THERMOSTABLE; ITS OPTIMAL
CC TEMPERATURE IS NEAR 65-70 DEGREES CELSIUS.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -!- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; M84339; AAA22265.1; -.
DR HSSP; P23904; 1AJK.
DR InterPro; IPR000757; Glyco_hydro_16.
DR Pfam; PF00722; Glyco_hydro_16; 1.
DR PRINTS; PR00737; GLHYDRLASE16.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 259 BETA-GLUCANASE.
FT ACT_SITE 142 142 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 146 146 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 259 AA; 29960 MW; A63C09F281FF5D13 CRC64;

Query Match 14.6%; Score 266.5; DB 1; Length 259;
Best Local Similarity 36.0%; Pred. No. 2.9e-12;
Matches 63; Conservative 24; Mismatches 75; Indels 13; Gaps 6;

```
Qy      26 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 85
           ||:::   || | :   || ||  || |   ||||| | |  |     || || |:||| |
Db      90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy      86 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 145
           ||:    | |  |   |    :| :: :   | ::|:||  ||  : : | |:|: |
Db     149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy     146 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQ 198
                 |:  | | :  |||       | | ||    |:: : :||   :||| |
Db     206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTPLQ 253
```

RESULT 3
GUB_BACSU
ID GUB_BACSU STANDARD; PRT; 242 AA.
AC P04957;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN BGLS OR BGL OR LICS OR N15B.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C120;
RX MEDLINE=84272222; PubMed=6087283;
RA Murphy N., McConnell D.J., Cantwell B.A.;
RT "The DNA sequence of the gene and genetic control sites for the
RT excreted B. subtilis enzyme beta-glucanase.";
RL Nucleic Acids Res. 12:5355-5367(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HL-25;
RA Tezuka H., Yuuki T., Yabuuchi S.;
RT "Construction of a beta-glucanase hyperproducing Bacillus subtilis
RT using the cloned beta-glucanase gene and a multi-copy plasmid.";
RL Agric. Biol. Chem. 53:2335-2339(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=97124196; PubMed=8969509;
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA Miwa Y., Fujita Y.;
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT containing the lic and cel loci, and creation of a 177 kb contig
RT covering the gnt-sacXY region.";
RL Microbiology 142:3113-3123(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95219081; PubMed=7704256;
RA Wolf M., Geczi A., Simon O., Borriss R.;
RT "Genes encoding xylan and beta-glucan hydrolysing enzymes in Bacillus
RT subtilis: characterization, mapping and construction of strains
RT deficient in lichenase, cellulase and xylanase.";
RL Microbiology 141:281-290(1995).
RN [5]
RP SEQUENCE OF 1-6 FROM N.A.
RC STRAIN=BR151;
RX MEDLINE=96178961; PubMed=8606172;
RA Schnetz K., Stuelke J., Gertz S., Krueger S., Krieg M., Hecker M.,

```
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00404; Dockerin_1; 2.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
DR   PROSITE; PS00018; EF_HAND; UNKNOWN_2.
DR   PROSITE; PS00448; CLOS_CELLULOSOME_RPT; 2.
KW   Hydrolase; Glycosidase; Signal; Repeat.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28    334       BETA-GLUCANASE.
FT   ACT_SITE    136    136       NUCLEOPHILE (BY SIMILARITY).
FT   ACT_SITE    140    140       PROTON DONOR (BY SIMILARITY).
FT   DOMAIN      252    269       PRO/THR-RICH (LINKER).
FT   DOMAIN      273    331       2 X 24 AA APPROXIMATE REPEATS.
FT   REPEAT      273    296       1.
FT   REPEAT      308    331       2.
FT   CONFLICT    304    334       QSVADVNRDGRIDSTDLTMLKRYLIRAIPSL ->
FT                                PQDGCGRHDRVVDSGSK (IN REF. 2 AND 3).
SQ   SEQUENCE   334 AA;  37897 MW;  0837564E9726F281 CRC64;

  Query Match             13.0%;  Score 237;  DB 1;  Length 334;
  Best Local Similarity   30.9%;  Pred. No. 5.1e-10;
  Matches   80;  Conservative   27;  Mismatches   94;  Indels   58;  Gaps   13;

Qy      28 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 81
           ::: |:  |   | :      || :| ||| |   | ||| | |   |   |  || |:|
Db      82 REYGGSYPYKSGEYRTKSFFGYGYYEVRMKAAKNVGIVSSFFTYTGPS---DNNPWDEID 138

Qy      82 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 141
           || |||:    | |      |    :|  | :   | | ||||| || |:|: : |||::
Db     139 IEFLGKDTTKVQFNWYKNGVGG---NEYLHNLGFDASQDFHTYGFEWRPDYIDFYVDGKK 195

Qy     142 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVY------- 192
           | :      |:  | | :  |||       |:|::|  : ||     :|| |
Db     196 VYR----GTRNIPVTPGKIMMNLWPGIGVDEWLGRYD-GRTPLQAEYEYVKYYPNGVPQD 250

Qy     193 ---------KYTP-----------GQGE-GGSDFT------LDWTDNFDTFDGSRWGKG 224
                     ||           | |    ||::      |   | |   | |
Db     251 NPTPTPTIAPSTPTNPNLPLKGDVNGDGHVNSSDYSLFKRYLLRVIDRFPVGDQS---VA 307

Qy     225 DWTFDGNRVDLTDKNIYSR 243
           |   || |:| ||  :  |
Db     308 DVNRDG-RIDSTDLTMLKR 325


RESULT 7
GUB_BACLI
ID   GUB_BACLI      STANDARD;      PRT;   243 AA.
AC   P27051;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE   (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN   BG1.
OS   Bacillus licheniformis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1402;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91224124; PubMed=2026156;
RA   Lloberas J., Perez-Pons J.A., Querol E.;
RT   "Molecular cloning, expression and nucleotide sequence of the
RT   endo-beta-1,3-1,4-D-glucanase gene from Bacillus licheniformis.
RT   Predictive structural analyses of the encoded polypeptide.";
RL   Eur. J. Biochem. 197:337-343(1991).
RN   [2]
RP   REVISIONS.
RA   Querol E.;
RL   Submitted (JUL-1991) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   MUTAGENESIS.
RX   MEDLINE=92362869; PubMed=1354172;
RA   Planas A., Juncosa M., Lloberas J., Querol E.;
RT   "Essential catalytic role of Glu134 in endo-beta-1,3-1,4-D-glucan 4-
RT   glucanohydrolase from B. licheniformis as determined by site-directed
RT   mutagenesis.";
RL   FEBS Lett. 308:141-145(1992).
RN   [4]
RP   MUTAGENESIS.
RX   MEDLINE=94237863; PubMed=8182059;
RA   Juncosa M., Pons J., Dot T., Querol E., Planas A.;
RT   "Identification of active site carboxylic residues in Bacillus
RT   licheniformis 1,3-1,4-beta-D-glucan 4-glucanohydrolase by
RT   site-directed mutagenesis.";
RL   J. Biol. Chem. 269:14530-14535(1994).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX   MEDLINE=96063718; PubMed=7589539;
RA   Hahn M., Pons J., Planas A., Querol E., Heinemann U.;
RT   "Crystal structure of Bacillus licheniformis 1,3-1,4-beta-D-glucan 4-
RT   glucanohydrolase at 1.8-A resolution.";
RL   FEBS Lett. 374:221-224(1995).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC       in beta-D-glucans containing 1,3- and 1,4-bonds.
CC   -!- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC       SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X57279; CAA40547.1; -.
DR   PIR; S15388; S15388.
DR   PDB; 1GBG; 07-DEC-95.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Hydrolase; Glycosidase; Signal; 3D-structure.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28    243       BETA-GLUCANASE.
FT   ACT_SITE    134    134       NUCLEOPHILE.
FT   ACT_SITE    138    138       PROTON DONOR.
FT   DISULFID     61     90
FT   MUTAGEN      51     51       D->N: 30% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN      89     89       D->N: 85% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN      92     92       E->Q: 3% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     105    105       E->Q: 50% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     133    133       D->N: 15% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     134    134       E->Q: 0.2% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     136    136       D->N: 0.5% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     138    138       E->Q: COMPLETE LOSS OF ACTIVITY.
FT   MUTAGEN     143    143       D->N: 65% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     160    160       E->Q: 15% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     168    168       D->N: 60% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     179    179       D->N: 80% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     190    190       D->N: 70% OF WILD-TYPE ACTIVITY.
FT   MUTAGEN     219    219       D->N: NO CHANGE IN ACTIVITY.
SQ   SEQUENCE   243 AA;  27435 MW;  651188D9AAD609A5 CRC64;

  Query Match             12.8%;  Score 233.5;  DB 1;  Length 243;
  Best Local Similarity   33.7%;  Pred. No. 6.1e-10;
  Matches   58;  Conservative   26;  Mismatches   77;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :|  || |   | ||| | |   :   || || |:||
Db      81 TSPSYNKFDCGENRSVQTYGYGLYEVNMKPAKNVGIVSSFFTYTGPT---DGTPWDEIDI 137

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |   |     ||   :   |  ::|||  :| || ::| |||| :
```

```
Db    138 EFLGKDTTKVQFNYYTNGVG---NHEKIVNLGFDAANSYHTYAFDWQPNSIKWYVDGQ-L 193

Qy    143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
          : |   |:    |   :  |||:      |:| ::    || : ::||:  |
Db    194 KHTATTQIPQTPGK--IMMNLWNGAGVDEWLGSYN-GVTPLSRSLHWVRYTK 242


RESULT 8
YG46_YEAST
ID   YG46_YEAST     STANDARD;      PRT;   507 AA.
AC   P53301;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Hypothetical 52.8 kDa protein in BUB1-HIP1 intergenic region.
GN   YGR189C OR G7553.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=97279231; PubMed=9133739;
RA   Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
RA   Nombela C.;
RT   "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
RT   of Saccharomyces cerevisiae chromosome VII.";
RL   Yeast 13:357-363(1997).
CC   -!- SIMILARITY: SOME, TO YEAST UTR2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z72974; CAA97215.1; -.
DR   EMBL; X99074; CAA67525.1; -.
DR   HSSP; P23904; 1AJK.
DR   SGD; S0003421; CRH1.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
KW   Hypothetical protein.
FT   DOMAIN       63     66       POLY-SER.
FT   DOMAIN      301    310       POLY-SER.
FT   DOMAIN      345    357       POLY-SER.
FT   DOMAIN      387    391       POLY-SER.
FT   DOMAIN      467    470       POLY-SER.
SQ   SEQUENCE   507 AA;  52757 MW;  7D7B61F57AEA942C CRC64;

  Query Match             12.6%;  Score 230;  DB 1;  Length 507;
  Best Local Similarity   26.9%;  Pred. No. 2.7e-09;
  Matches   87;  Conservative   57;  Mismatches  134;  Indels   46;  Gaps   13;

Qy      6 TAVKSALAVAAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLY 65
          | :| |  :   :  |:  : || :    | :   : ||| |  :| |  :| ||| :|
Db     70 TDLKHAGEIKYGSDGLSMTL-AKRYDNPSLKSNFYIMYGKLEVILKAANGTGIVSSFYLQ 128

Qy     66 QNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYG 125
            :  :        |:||| :| :   ||||  :          : | |    |: || |
Db    129 SDDLD--------EIDIEWVGGDNTQFQSNFFSKGDTTTYDRGEFHGVDTPTDK-FHNYT 179

Qy    126 LEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQG-------LRFNLWSS---ESAA----W 171
          |:|  :   | :||: ||      | : | ::|       |   :|:    ::||    |
Db    180 LDWAMDKTTWYLDGESVR------VLSNTSSEGYPQSPMYLMMGIWAGGDPDNAAGTIEW 233

Qy    172 V-GQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDG 230
            |: : :  |   :|  | |  |: |:     | :  | :: :  \|| :|: |   :
Db    234 AGGETNYNDAPFTMYIEKVIVTDYSTGKKYTYGDQSGSW-ESIEADGGSIYGRYDQAQE- 291

Qy    231 NRVDLTDKNIYSRDGMLILALTRKG--QESFNGQVPRDDEPAPQSSSSAPASS--SSVPA 286
                |  : :  | :  : |       | :  |         ||:|:  ||  ||  :
Db    292 ------DFAVLANGGSISSSSTSSSTVSSSASSTVSSSVSSTVSSSASSTVSSSVSSTVS 345

Qy    287 SSSSVPASSSSAFVPPSSSSATNA 310
          |||||  ||||::   ||||:||::
Db    346 SSSSVSSSSSTS---PSSSTATSS 366


RESULT 9
GUB_PAEMA
ID   GUB_PAEMA      STANDARD;      PRT;   237 AA.
AC   P23904;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE   (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
OS   Paenibacillus macerans (Bacillus macerans).
OC   Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
OX   NCBI_TaxID=44252;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91109712; PubMed=2274030;
RA   Borriss R., Buettner K., Maentsaelae P.;
RT   "Structure of the beta-1,3-1,4-glucanase gene of Bacillus macerans:
RT   homologies to other beta-glucanases.";
RL   Mol. Gen. Genet. 222:278-283(1990).
RN   [2]
RP   ACTIVE SITE.
RX   MEDLINE=93094208; PubMed=1360982;
RA   Hoej P.B., Condron R., Traeger J.C., McAuliffe J.C., Stone B.A.;
RT   "Identification of glutamic acid 105 at the active site of Bacillus
RT   amyloliquefaciens 1,3-1,4-beta-D-glucan 4-glucanohydrolase using
RT   epoxide-based inhibitors.";
RL   J. Biol. Chem. 267:25059-25066(1992).
RN   [3]
RP   MUTAGENESIS OF GLU-128.
RA   Olsen O.;
RL   Thesis (1990), University of Aarhus, Denmark.
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=93281743; PubMed=8099449;
RA   Keitel T., Simon O., Borriss R., Heinemann U.;
RT   "Molecular and active-site structure of a Bacillus 1,3-1,4-beta-
RT   glucanase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:5287-5291(1993).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX   MEDLINE=96028129; PubMed=7588726;
RA   Hahn M., Keitel T., Heinemann U.;
RT   "Crystal and molecular structure at 0.16-nm resolution of the hybrid
RT   Bacillus endo-1,3-1,4-beta-D-glucan 4-glucanohydrolase H(A16-M).";
RL   Eur. J. Biochem. 232:849-858(1995).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC       in beta-D-glucans containing 1,3- and 1,4-bonds.
CC   -!- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC       SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X55959; CAA39426.1; -.
DR   PIR; S11927; S11927.
DR   PDB; 2AYH; 31-MAR-95.
DR   PDB; 1BYH; 31-OCT-93.
DR   PDB; 1CPM; 22-JUN-94.
```

```
RA   Rak B.;
RT   "LicT, a Bacillus subtilis transcriptional antiterminator protein of
RT   the BglG family.";
RL   J. Bacteriol. 178:1971-1979(1996).
RN   [6]
RP   SEQUENCE OF 29-63.
RA   Yuuki T., Tezuka H., Yabuuchi S.;
RT   "Purification and some properties of two enzymes from a beta-glucanase
RT   hyperproducing strain, Bacillus subtilis HL-25.";
RL   Agric. Biol. Chem. 53:2341-2346(1989).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC       in beta-D-glucans containing 1,3- and 1,4-bonds.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC       SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC   -!- MISCELLANEOUS: THE SEQUENCE OF STRAIN 168 IS SHOWN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC - modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X00754; CAA25328.1; -.
DR   EMBL; D00518; BAA00405.1; -.
DR   EMBL; D83026; BAA11697.1; -.
DR   EMBL; Z46862; CAA86922.1; -.
DR   EMBL; Z28340; CAA82195.1; -.
DR   EMBL; Z99124; CAB15943.1; -.
DR   PIR; A22914; LXBS.
DR   PIR; JU0110; JU0110.
DR   HSSP; P27051; 1GBG.
DR   SubtiList; BG10476; bglS.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Hydrolase; Glycosidase; Signal; Complete proteome.
FT   SIGNAL       1     28
FT   CHAIN       29    242       BETA-GLUCANASE.
FT   MOD_RES     29     29       PYRROLIDONE CARBOXYLIC ACID.
FT   ACT_SITE   133    133       NUCLEOPHILE (BY SIMILARITY).
FT   ACT_SITE   137    137       PROTON DONOR (BY SIMILARITY).
FT   DISULFID    60     89       BY SIMILARITY.
FT   VARIANT     24     24       A -> S (IN STRAIN HL-25).
FT   VARIANT     83     83       A -> S (IN STRAIN HL-25).
FT   VARIANT    204    204       P -> L (IN STRAIN C120).
SQ   SEQUENCE   242 AA;  27268 MW;  45958DEA70F22B29 CRC64;

  Query Match             13.6%;  Score 248;  DB 1;  Length 242;
  Best Local Similarity   36.0%;  Pred. No. 5.6e-11;
  Matches   63;  Conservative   25;  Mismatches   75;  Indels   12;  Gaps    7;

Qy      20 ALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVE 79
           |||:    | |    |  :::   || :| ||| |  :| ||| | |    :   || || |
Db      78 ALTSPAYNK-FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDE 133

Qy      80 VDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 139
           :||| |||:    | |  |  ||     ||   :   |  |:||  :| || ::| |||
Db     134 IDIEFLGKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDG 190

Qy     140 QEVRKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           | :: |   |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     191 Q-LKHTATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 4
GUB_BACAM
ID   GUB_BACAM      STANDARD;      PRT;   239 AA.
AC   P07980;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE   (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN   BGLA.
OS   Bacillus amyloliquefaciens.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1390;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BE 20/78;
RX   MEDLINE=87192007; PubMed=3106158;
RA   Hofemeister J., Kurtz A., Borriss R., Knowles J.;
RT   "The beta-glucanase gene from Bacillus amyloliquefaciens shows
RT   extensive homology with that of Bacillus subtilis.";
RL   Gene 49:177-187(1985).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC       in beta-D-glucans containing 1,3- and 1,4-bonds.
CC   -!- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC       SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M15674; AAA87323.1; -.
DR   PIR; A29091; A29091.
DR   HSSP; P27051; 1GBG.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Hydrolase; Glycosidase; Signal.
FT   SIGNAL       1     25
FT   CHAIN       26    239       BETA-GLUCANASE.
FT   ACT_SITE   134    134       NUCLEOPHILE (BY SIMILARITY).
FT   DISULFID    57     86       BY SIMILARITY.
SQ   SEQUENCE   239 AA;  26928 MW;  A76A64268A7AAA0B CRC64;

  Query Match             13.3%;  Score 242.5;  DB 1;  Length 239;
  Best Local Similarity   34.3%;  Pred. No. 1.4e-10;
  Matches   59;  Conservative   27;  Mismatches   75;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |  |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db      77 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 133

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :
Db     134 EFLGKDTTKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 189

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           : |   |:    |   :  |||:      |:| ::    |::   :|::  |
Db     190 KHTATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYDWMRYRK 238


RESULT 5
XYND_RUMFL
ID   XYND_RUMFL     STANDARD;      PRT;   802 AA.
AC   Q53317;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Xylanase/beta-glucanase precursor [Includes: Endo-1,4-beta-xylanase
DE   (EC 3.2.1.8) (Xylanase); Endo-beta-1,3-1,4 glucanase (EC 3.2.1.73)
DE   (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase)].
GN   XYND.
```

```
OS   Ruminococcus flavefaciens.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC   Ruminococcus.
OX   NCBI_TaxID=1265;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=17;
RX   MEDLINE=93259938; PubMed=8491715;
RA   Flint H.J., Martin J., McPherson C.A., Daniel A.S., Zhang J.-X.;
RT   "A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-
RT   glucanase domains, encoded by the xynD gene of Ruminococcus
RT   flavefaciens.";
RL   J. Bacteriol. 175:2943-2951(1993).
CC   -!- FUNCTION: CONTAINS TWO CATALYTIC DOMAINS WITH XYLANASE AND ENDO-
CC       BETA-1,3-1,4 GLUCANASE ACTIVITIES.
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC       linkages in xylans.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC       in beta-D-glucans containing 1,3- and 1,4-bonds.
CC   -!- PATHWAY: Xylan degradation.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
CC       G (FAMILY 11 OF GLYCOSYL HYDROLASES).
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO FAMILY 16 OF
CC       GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S61204; AAB26620.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR003305; CBM_CenC.
DR   InterPro; IPR001137; GH_11.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00457; Glyco_hydro_11; 1.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   Pfam; PF02018; CBM_4_9; 1.
DR   PRINTS; PR00911; GLHYDRLASE11.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS00776; GLYCOSYL_HYDROL_F11_1; 1.
DR   PROSITE; PS00777; GLYCOSYL_HYDROL_F11_2; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Xylan degradation; Hydrolase; Glycosidase; Signal;
KW   Multifunctional enzyme.
FT   SIGNAL        1     31       POTENTIAL.
FT   CHAIN        32    802       XYLANASE/BETA-GLUCANASE.
FT   DOMAIN       32    244       A (XYLANASE).
FT   DOMAIN      245    523       B.
FT   DOMAIN      524    555       LINKER.
FT   DOMAIN      556    802       C (BETA-GLUCANASE).
FT   ACT_SITE    124    124       NUCLEOPHILE (BY SIMILARITY).
FT   ACT_SITE    226    226       PROTON DONOR (BY SIMILARITY).
FT   ACT_SITE    684    684       NUCLEOPHILE (BY SIMILARITY).
FT   DOMAIN      524    529       POLY-THR.
FT   DOMAIN      532    543       POLY-THR.
FT   DOMAIN      546    553       POLY-THR.
SQ   SEQUENCE   802 AA;  89091 MW;  2880A689647284AF CRC64;

  Query Match           13.3%;  Score 242.5;  DB 1;  Length 802;
  Best Local Similarity  34.9%;  Pred. No. 6e-10;
  Matches   58;  Conservative   21;  Mismatches   74;  Indels   13;  Gaps    6;

Qy      30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
           :|| |  |     || :|  |:     | ||| | |   |   | \|| |:|||:||||
Db     634 YSGGEFRTNNFYHYGYYECSMQAMKNDGVVSSFFTYTGPS---DDNPWDEIDIEILGKNT 690

Qy      90 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 149
             | |  |   |     || : :   : :|:|||| :| |||: | |||:|| :
Db     691 TQVQFNYYTNGQGKH---EKLYDLGFDSSEAYHTYGFDWQPNYIAWYVDGREVYRA---- 743
```

```
Qy     150 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           ::  | | :  | |   :   |:  |:  : ||     ||   |
Db     744 TQDIPKTPGKIMMNAWPGLTVDDWLKAFN-GRTPLTAHYQWVTYNK 788


RESULT 6
GUB_CLOTM
ID   GUB_CLOTM      STANDARD;      PRT;   334 AA.
AC   P29716; P37074;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Beta-glucanase precursor (EC 3.2.1.73). (Endo-beta-1,3-1,4 glucanase)
DE   (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase) (Laminarinase).
GN   LICB OR LAM1.
OS   Clostridium thermocellum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1515;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 27405 / DSM 1237;
RX   MEDLINE=92155194; PubMed=1740123;
RA   Schimming S., Schwarz W.H., Staudenbauer W.L.;
RT   "Structure of the Clostridium thermocellum gene licB and the encoded
RT   beta-1,3-1,4-glucanase. A catalytic region homologous to Bacillus
RT   lichenases joined to the reiterated domain of clostridial
RT   cellulases.";
RL   Eur. J. Biochem. 204:13-19(1992).
RN   [2]
RP   PRELIMINARY SEQUENCE FROM N.A.
RC   STRAIN=F7;
RX   MEDLINE=92095946; PubMed=1755832;
RA   Zverlov V.V., Laptev D.A., Tishkov V.I., Velikodvorskaja G.A.;
RT   "Nucleotide sequence of the Clostridium thermocellum laminarinase
RT   gene.";
RL   Biochem. Biophys. Res. Commun. 181:507-512(1991).
RN   [3]
RP   REVISIONS.
RA   Zverlov V.V.;
RL   Submitted (AUG-1993) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   CHARACTERIZATION.
RC   STRAIN=F7;
RA   Zverlov V.V., Velikodvorskaja G.A.;
RT   "Cloning the Clostridium thermocellum thermostable laminarinase gene
RT   in Escherichia coli; the properties of the enzyme thus produced.";
RL   Biotechnol. Lett. 12:811-816(1990).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC       in beta-D-glucans containing 1,3- and 1,4-bonds.
CC   -!- SUBUNIT: MAY FORM PART OF A MULTIENZYME COMPLEX (CELLULOSOME).
CC   -!- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
CC       WELL AS IN OTHER C.THERMOCELLUM CELLULOSOME ENZYMES. THIS DOMAIN
CC       MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X63355; CAA44959.1; -.
DR   EMBL; X58392; CAA41281.1; -.
DR   PIR; S23498; S23498.
DR   PIR; JS0611; JS0611.
DR   PIR; S18726; S18726.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR002105; Dockerin_1.
DR   InterPro; IPR002048; EF-hand.
```

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:50:42 ; Search time 125.774 Seconds
(without alignments)
1368.672 Million cell updates/sec

Title: US-09-654-652A-2
Perfect score: 1439
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....PNSSSVDKLAAALEHHHHHH 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1439	100.0	267	20	US-09-654-652A-2	Sequence 2, Appli
2	1340	93.1	349	20	US-09-654-652A-3	Sequence 3, Appli
3	1340	93.1	349	21	US-09-791-537-30441	Sequence 30441, A
4	1333	92.6	248	20	US-09-654-652A-1	Sequence 1, Appli
5	272.5	18.9	214	21	US-09-791-537-145019	Sequence 145019,
6	266.5	18.5	259	21	US-09-791-537-106873	Sequence 106873,
7	265.5	18.5	252	21	US-09-791-537-32335	Sequence 32335, A
8	257.5	17.9	308	18	US-09-463-862A-1	Sequence 1, Appli
9	256.5	17.8	208	21	US-09-791-537-6910	Sequence 6910, Ap
10	250.5	17.4	394	21	US-09-791-537-132244	Sequence 132244,
11	247.5	17.2	214	21	US-09-791-537-117361	Sequence 117361,
12	247.5	17.2	276	21	US-09-791-537-31771	Sequence 31771, A
13	246.5	17.1	242	21	US-09-791-537-27063	Sequence 27063, A
14	245.5	17.1	242	21	US-09-791-537-73450	Sequence 73450, A
15	245.5	17.1	242	21	US-09-791-537-79889	Sequence 79889, A
16	245	17.0	239	21	US-09-791-537-103347	Sequence 103347,
17	244	17.0	239	21	US-09-791-537-103324	Sequence 103324,
18	242.5	16.9	802	21	US-09-791-537-50107	Sequence 50107, A
19	242.5	16.9	802	21	US-09-791-537-116925	Sequence 116925,
20	238.5	16.6	239	21	US-09-791-537-30436	Sequence 30436, A
21	237	16.5	334	21	US-09-791-537-48909	Sequence 48909, A
22	235.5	16.4	214	21	US-09-791-537-88552	Sequence 88552, A
23	234.5	16.3	239	21	US-09-791-537-18679	Sequence 18679, A
24	233	16.2	239	21	US-09-791-537-103817	Sequence 103817,
25	231.5	16.1	320	21	US-09-791-537-67048	Sequence 67048, A
26	230	16.0	237	21	US-09-791-537-72848	Sequence 72848, A
27	229.5	15.9	243	21	US-09-791-537-30440	Sequence 30440, A
28	226	15.7	212	21	US-09-791-537-88873	Sequence 88873, A
29	226	15.7	214	21	US-09-791-537-76270	Sequence 76270, A
30	226	15.7	237	21	US-09-791-537-103320	Sequence 103320,
31	226	15.7	237	21	US-09-791-537-106876	Sequence 106876,
32	226	15.7	239	21	US-09-791-537-103345	Sequence 103345,
33	226	15.7	239	21	US-09-791-537-103350	Sequence 103350,
34	226	15.7	239	21	US-09-791-537-103355	Sequence 103355,
35	226	15.7	239	21	US-09-791-537-103357	Sequence 103357,
36	226	15.7	239	21	US-09-791-537-103814	Sequence 103814,
37	226	15.7	239	21	US-09-791-537-103819	Sequence 103819,
38	226	15.7	240	21	US-09-791-537-64827	Sequence 64827, A
39	225	15.6	238	21	US-09-791-537-79887	Sequence 79887, A
40	222.5	15.5	242	21	US-09-791-537-37095	Sequence 37095, A
41	222	15.4	237	21	US-09-791-537-3485	Sequence 3485, Ap
42	214	14.9	160	21	US-09-706-729-14	Sequence 14, Appl
43	209.5	14.6	504	27	US-60-385-568-329	Sequence 329, App
44	209.5	14.6	526	25	US-10-179-131-6380	Sequence 6380, Ap
45	208.5	14.5	526	16	US-09-248-796-14807	Sequence 14807, A

ALIGNMENTS

RESULT 1
US-09-654-652A-2
; Sequence 2, Application US/09654652A
; GENERAL INFORMATION:
; APPLICANT: SHYUR, LIE-FEN
; APPLICANT: CHEN, JUI-LIN
; APPLICANT: YANG, NING-SUN
; TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1,
; TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
; TITLE OF INVENTION: THERMO-TOLERANCE
; FILE REFERENCE: 4910-3
; CURRENT APPLICATION NUMBER: US/09/654,652A
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme
; OTHER INFORMATION: with enhanced activity and thermal stability
US-09-654-652A-2

Query Match 100.0%; Score 1439; DB 20; Length 267;
Best Local Similarity 100.0%; Pred. No. 5.1e-138;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60

Qy     61 VLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120

Qy    121 KTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 KTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS 180

Qy    181 DFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQV 240

Qy    241 PRDDEPAPNSSSVDKLAAALEHHHHHH 267
          ||||||||||||||||||||||||||
Db    241 PRDDEPAPNSSSVDKLAAALEHHHHHH 267


RESULT 2
US-09-654-652A-3
; Sequence 3, Application US/09654652A
; GENERAL INFORMATION:
;  APPLICANT: SHYUR, LIE-FEN
;  APPLICANT:  CHEN, JUI-LIN
;  APPLICANT:  YANG, NING-SUN
;  TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1,
;  TITLE OF INVENTION:  4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
;  TITLE OF INVENTION:  THERMO-TOLERANCE
;  FILE REFERENCE: 4910-8
;  CURRENT APPLICATION NUMBER: US/09/654,652A
;  CURRENT FILING DATE:  2000-09-05
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Fibrobacter succinogenes
US-09-654-652A-3

  Query Match             93.1%;  Score 1340;  DB 20;  Length 349;
  Best Local Similarity   99.6%;  Pred. No. 1e-127;
  Matches  250;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84

Qy     62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144

Qy    122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204

Qy    182 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 264

Qy    242 RDDEPAPNSSS 252
          ||||||| |||
Db    265 RDDEPAPQSSS 275


RESULT 3
US-09-791-537-30441
; Sequence 30441, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30441
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Fibrobacter succinogenes
US-09-791-537-30441

  Query Match             93.1%;  Score 1340;  DB 21;  Length 349;
  Best Local Similarity   99.6%;  Pred. No. 1e-127;
  Matches  250;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84

Qy     62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144

Qy    122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204

Qy    182 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 264

Qy    242 RDDEPAPNSSS 252
          ||||||| |||
Db    265 RDDEPAPQSSS 275


RESULT 4
US-09-654-652A-1
; Sequence 1, Application US/09654652A
; GENERAL INFORMATION:
;  APPLICANT: SHYUR, LIE-FEN
;  APPLICANT:  CHEN, JUI-LIN
;  APPLICANT:  YANG, NING-SUN
;  TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1,
;  TITLE OF INVENTION:  4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
;  TITLE OF INVENTION:  THERMO-TOLERANCE
;  FILE REFERENCE: 4910-8
;  CURRENT APPLICATION NUMBER: US/09/654,652A
;  CURRENT FILING DATE:  2000-09-05
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 248
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme
;   OTHER INFORMATION: with enhanced activity and thermal stability
US-09-654-652A-1

  Query Match             92.6%;  Score 1333;  DB 20;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-127;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60

Qy     61 VLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120
```

```
Qy    121 KTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 KTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS 180

Qy    181 DFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQV 240

Qy    241 PRDDEPAP 248
          ||||||||
Db    241 PRDDEPAP 248


RESULT 5
US-09-791-537-145019
; Sequence 145019, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 145019
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: pdb 1CPM
US-09-791-537-145019

  Query Match             18.9%;  Score 272.5;  DB 21;  Length 214;
  Best Local Similarity   33.8%;  Pred. No. 7.7e-19;
  Matches   77;  Conservative   33;  Mismatches   93;  Indels   25;  Gaps   12;

Qy      7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
          |  ||  :     || :|  || |  :| ||| | |       | |  | |:||| |||:
Db      1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 57

Qy     67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
          | |  |   |     ||  ::   | : |||| :| | |::| |||  :: |
Db     58 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 110

Qy    127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTL 184
           :|:  | | :  |||:      |:| :: :  ||:   :||   |||  |  ||| |
Db    111 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYTSNQ-TGGSFF-- 163

Qy    185 DWTDNFDTFDGSRWGKGDWTFDGNRVDLT--DKNI-YSRDGMLILALT 229
          : |::::   | | |   :|   : |    |: :: || | | ||
Db    164 ---EPFNSYNSGTWEKADGYSNGGVFNCTWRANNVNFTNDGKLKLGLT 208


RESULT 6
US-09-791-537-106873
; Sequence 106873, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 106873
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Brevibacillus brevis
US-09-791-537-106873

  Query Match             18.5%;  Score 266.5;  DB 21;  Length 259;
  Best Local Similarity   36.0%;  Pred. No. 4.1e-18;
  Matches   63;  Conservative   24;  Mismatches   75;  Indels   13;  Gaps    6;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          ||:::   || | :   || ||  || |   ||||| | |  |    || || |:||| |
Db     90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
          ||:    | |  |   |    :| :: :   | ::|:||  ||  : : | |:|: |
Db    149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy    123 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQ 175
              |:  | | :  |||       | | ||    |:: : :||   :||| |
Db    206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTPLQ 253


RESULT 7
US-09-791-537-32335
; Sequence 32335, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32335
;   LENGTH: 252
;   TYPE: PRT
;   ORGANISM: Brevibacillus brevis
US-09-791-537-32335

  Query Match             18.5%;  Score 265.5;  DB 21;  Length 252;
  Best Local Similarity   35.8%;  Pred. No. 5e-18;
  Matches   62;  Conservative   24;  Mismatches   74;  Indels   13;  Gaps    6;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          ||:::   || | :   || ||  || |   ||||| | |  |    || || |:||| |
Db     90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
          ||:    | |  |   |    :| :: :   | ::|:||  ||  : : | |:|: |
Db    149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy    123 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
              |:  | | :  |||       | | ||    |:: : :||   :||| |
Db    206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTP 251


RESULT 8
US-09-463-862A-1
; Sequence 1, Application US/09463862A
; GENERAL INFORMATION:
;  APPLICANT: Hillen, Wolfgang
;  APPLICANT:  Maurer, Karl-Heinz
;  TITLE OF INVENTION: New Beta-Glucanase from a Bacillus
;  FILE REFERENCE: H2081PCT/US
;  CURRENT APPLICATION NUMBER: US/09/463,862A
;  CURRENT FILING DATE:  2000-05-01
;  PRIOR APPLICATION NUMBER: DE 197 32 751.6
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: PCT/EP98/04564
;  PRIOR FILING DATE: 1998-07-21
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
```

```
;   LENGTH: 308
;   TYPE: PRT
;   ORGANISM: Bacillus alkalophilus DSM 9956
US-09-463-862A-1

  Query Match             17.9%;  Score 257.5;  DB 18;  Length 308;
  Best Local Similarity   36.5%;  Pred. No. 4.4e-17;
  Matches   62;  Conservative   22;  Mismatches   73;  Indels   13;  Gaps    6;

Qy       6 DFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN 65
           ::   || | :  ||| ||  || | ::|||||:| |  |    |  || |:||| |||:
Db      96 EYKAGELRTNQFYQYGLFEVNMKPAKSTGTVSSLFTY-TGPWDWDNDPWDEIDIEFLGKD 154

Qy      66 PGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGG 125
             | |  |   |    :| :| :   | ::|:||  || |  : | |:|: |
Db     155 TTRVQFNYFTNGVG---NNEHYHELGFDASESFNTYAFEWRPESISWYVNGELVYTA--- 208

Qy     126 QVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             |:  | | :  |||       | | ||    |:    :||   :|||
Db     209 -TENIPQTPQKIMMNLWPGIGVDGWTGVFDGEDTPVVTEYDWV---RYTP 254


RESULT 9
US-09-791-537-6910
; Sequence 6910, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6910
;   LENGTH: 208
;   TYPE: PRT
;   ORGANISM: pdb 1CPN
US-09-791-537-6910

  Query Match             17.8%;  Score 256.5;  DB 21;  Length 208;
  Best Local Similarity   33.8%;  Pred. No. 3.2e-17;
  Matches   77;  Conservative   29;  Mismatches   95;  Indels   27;  Gaps   13;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |  ||  :     || :|  || |  :| ||| | |      | |  | |:||| |||:
Db       1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 57

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db      58 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 110

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTL 184
           :|:  | | :  |||:      |:| :: :  ||:   :||   |||      || |
Db     111 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT----SNGSVF-- 160

Qy     185 DWTDNFDTFDGSRWGKGDWTFDGNRVDLT--DKNI-YSRDGMLILALT 229
             |      |: | | | |   :|   : |    |: :: || | | ||
Db     161 -WEPK-SYFNPSTWEKADGYSNGGVFNCTWRANNVNFTNDGKLKLGLT 206


RESULT 10
US-09-791-537-132244
; Sequence 132244, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132244
;   LENGTH: 394
;   TYPE: PRT
;   ORGANISM: pdb 1AXKA
US-09-791-537-132244

  Query Match             17.4%;  Score 250.5;  DB 21;  Length 394;
  Best Local Similarity   29.6%;  Pred. No. 3.3e-16;
  Matches   68;  Conservative   33;  Mismatches   90;  Indels   39;  Gaps   10;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |  ||  :     || :|  || |  :| ||| | |      | |  | |:||| |||:
Db       1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 57

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db      58 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 110

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKV---------YKYTPGQ 175
           :|:  | | :  |||:      |:| :: :  ||:   :|||            :| |
Db     111 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWVKYTSNASTDYWQNWTDGG 169

Qy     176 G------EGGSDFTLDWTDNFDTFDGSRWGKGD----------WTFDGN 208
           |      | :::::|::  :   |  |  |              |  :||
Db     170 GIVNAVNGSGGNYSVNWSNTGNFVVGKGWTTGSPFRTINYNAGVWAPNGN 219


RESULT 11
US-09-791-537-117361
; Sequence 117361, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 117361
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: pdb 1AJKA
US-09-791-537-117361

  Query Match             17.2%;  Score 247.5;  DB 21;  Length 214;
  Best Local Similarity   32.9%;  Pred. No. 2.8e-16;
  Matches   69;  Conservative   32;  Mismatches   84;  Indels   25;  Gaps   12;

Qy      33 SGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSP 92
           :| ||| | |      | |  | |:||| |||:    | |  |   |     ||  ::
Db       2 TGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDTTKVQFNYYTNGVGGH---EKVISLGF 55

Qy      93 AADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVG 150
           | : ||||  :| | |::| |||  :: |     :|:  | | :  |||:      |:|
Db      56 DASKGFHTYAFDWQPGYIKWYVDG-VLKHT---ATANIPSTPGKIMMNLWNGTGVDDWLG 111

Qy     151 QFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRV 210
           :: :  ||:   :||   |||  |  ||| |     : |::::   | | |   :|
Db     112 SYNGAN-PLYAEYDWV---KYTSNQ-TGGSFF-----EPFNSYNSGTWEKADGYSNGGVF 161

Qy     211 DLT--DKNI-YSRDGMLILALTRKGQESFN 237
           : |    |: :: || | | ||      |:
Db     162 NCTWRANNVNFTNDGKLKLGLTSSAYNKFD 191


RESULT 12
```

```
US-09-791-537-31771
; Sequence 31771, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31771
;   LENGTH: 276
;   TYPE: PRT
;   ORGANISM: Bacillus sp
US-09-791-537-31771

  Query Match             17.2%;  Score 247.5;  DB 21;  Length 276;
  Best Local Similarity   36.1%;  Pred. No. 4e-16;
  Matches   61;  Conservative   20;  Mismatches   75;  Indels   13;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db      96 YKAGELRTNDFYHYGLFEVSMKPAKSTGTVSSFFTY-TGPWDWENDPWDEIDIEFLGKDT 154

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
Db     155 TKIQFNYFTNGVGG---NEHYHELGFDAADDFNTYAFEWRPESIRWFVNGELVHTA---- 207

Qy     127 VSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
Db     208 TENIPQTPQKIMMNLWPGIGVDGWTGRFNGEDTPVVTQYDWV---KYTP 253


RESULT 13
US-09-791-537-27063
; Sequence 27063, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27063
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-791-537-27063

  Query Match             17.1%;  Score 246.5;  DB 21;  Length 242;
  Best Local Similarity   35.5%;  Pred. No. 4.2e-16;
  Matches   60;  Conservative   24;  Mismatches   74;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db      83 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFL 139

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
Db     140 GKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 195

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
Db     196 ATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 14
US-09-791-537-73450
; Sequence 73450, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73450
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-791-537-73450

  Query Match             17.1%;  Score 245.5;  DB 21;  Length 242;
  Best Local Similarity   35.8%;  Pred. No. 5.3e-16;
  Matches   59;  Conservative   24;  Mismatches   71;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db      87 FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFLGKDT 143

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
Db     144 TKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATNQ 199

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
Db     200 IPTTLGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 15
US-09-791-537-79889
; Sequence 79889, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 79889
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-791-537-79889

  Query Match             17.1%;  Score 245.5;  DB 21;  Length 242;
  Best Local Similarity   35.8%;  Pred. No. 5.3e-16;
  Matches   59;  Conservative   24;  Mismatches   71;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db      87 FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFLGKDT 143

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
Db     144 TKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATNQ 199

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
Db     200 IPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241
```

Search completed: January 9, 2003, 12:13:28
Job time : 127.774 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:13:37 ; Search time 31.103 Seconds
(without alignments)
2312.009 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKKTAVKSALAVAAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	268.5	14.7	851	16	Q9K7X5	Q9k7x5 bacillus ha
2	254.5	14.0	256	2	Q9APD8	Q9apd8 bacillus ci
3	247.5	13.6	276	2	Q45648	Q45648 bacillus sp
4	246.5	13.5	214	2	Q93GE8	Q93ge8 uncultured
5	245.5	13.5	214	2	Q93GE7	Q93ge7 uncultured
6	242.5	13.3	802	2	Q9S310	Q9s310 ruminococcu
7	239.5	13.1	214	2	Q93GE6	Q93ge6 uncultured
8	238.5	13.1	239	2	Q45691	Q45691 bacillus su
9	236.5	13.0	242	2	Q8RMP0	Q8rmp0 bacillus su
10	232.5	12.7	212	2	Q9FDC9	Q9fdc9 paenibacill
11	230	12.6	237	2	O07856	O07856 streptococc
12	223	12.2	205	2	Q93UI1	Q93ui1 uncultured
13	221	12.1	246	16	Q97FD3	Q97fd3 clostridium
14	206	11.3	289	16	Q8U8N5	Q8u8n5 agrobacteri
15	200	11.0	293	16	Q98C78	Q98c78 rhizobium l
16	190.5	10.4	245	3	O14412	O14412 orpinomyces
17	190.5	10.4	302	16	Q9K7X6	Q9k7x6 bacillus ha
18	185	10.1	282	10	Q9FI31	Q9fi31 arabidopsis
19	178	9.8	302	3	O42800	O42800 aspergillus
20	173	9.5	282	10	O80803	O80803 arabidopsis
21	173	9.5	642	16	Q9WXN1	Q9wxn1 thermotoga
22	172.5	9.5	277	10	Q9M0D1	Q9m0d1 arabidopsis
23	171	9.4	646	2	Q60039	Q60039 thermotoga
24	170.5	9.3	282	10	Q9M0D2	Q9m0d2 arabidopsis
25	169.5	9.3	277	10	Q94A49	Q94a49 arabidopsis
26	164	9.0	286	10	Q38910	Q38910 arabidopsis
27	161	8.8	284	10	Q9FKL8	Q9fkl8 arabidopsis
28	160	8.8	286	10	Q8W4M6	Q8w4m6 arabidopsis
29	159.5	8.7	287	10	Q9ZSU4	Q9zsu4 arabidopsis
30	159.5	8.7	306	2	Q51333	Q51333 oerskovia x
31	157.5	8.6	92	3	Q9HGU1	Q9hgu1 aspergillus
32	157.5	8.6	94	3	Q9P420	Q9p420 trichoderma
33	157.5	8.6	269	10	Q39148	Q39148 arabidopsis
34	156	8.6	292	10	Q9ZRV1	Q9zrv1 fagus sylva
35	155	8.5	845	2	Q9KWF3	Q9kwf3 clostridium
36	155	8.5	845	16	Q8XNF8	Q8xnf8 clostridium
37	154.5	8.5	277	10	Q38907	Q38907 arabidopsis
38	154.5	8.5	284	10	Q9SEB0	Q9seb0 arabidopsis
39	153	8.4	163	2	Q93UI2	Q93ui2 uncultured
40	149	8.2	289	10	P93669	P93669 hordeum vul
41	146.5	8.0	284	10	Q38857	Q38857 arabidopsis
42	145.5	8.0	285	10	Q9FKL9	Q9fkl9 arabidopsis
43	145.5	8.0	310	10	Q9XIJ7	Q9xij7 arabidopsis
44	145	7.9	280	10	Q949I0	Q949i0 festuca pra
45	145	7.9	305	10	Q9ZV40	Q9zv40 arabidopsis

ALIGNMENTS

RESULT 1
Q9K7X5
ID Q9K7X5 PRELIMINARY; PRT; 851 AA.
AC Q9K7X5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Endo-beta-1,3-1,4 glucanase (licheninase) (EC 3.2.1.73).
GN BGLS OR BH3232.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001518; BAB06951.1; -.
DR HSSP; P23904; 1AJK.
DR InterPro; IPR000757; Glyco_hydro_16.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00722; Glyco_hydro_16; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR PRINTS; PR00737; GLHYDRLASE16.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 3.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 851 AA; 98142 MW; C1C4F0F158400285 CRC64;

Query Match 14.7%; Score 268.5; DB 16; Length 851;
Best Local Similarity 29.4%; Pred. No. 7.2e-10;
Matches 83; Conservative 32; Mismatches 112; Indels 55; Gaps 12;

Qy 21 LTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEV 80

```
          |  :  |  :   || | :  ||| ||  || |   |||| | |  |    || || |:
Db     93 LENDEQAPPYKAGELRTNDFYQYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEI 151

Qy     81 DIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQ 140
          ||| |||:    | |  |    |    :| :: :   | ::|:||  ||  : : | |:|:
Db    152 DIEFLGKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGE 208

Qy    141 EVRKTEGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP-- 196
           |         |:  | | :  |||       | | ||    ||     ||   :|||
Db    209 AVHTA----TENIPQTPQKIMMNLWPGVGVDEWTGVFDGDNTPLHADYEWV---RYTPLE 261

Qy    197 --GQGEGGSDFTLDWT----------------------DNFDTFDGSRWG-KGDWTFDG 230
             : | | |   :                        : |:||:   |     || :|
Db    262 ALDEEESGDDNEEEPVEEVEEEEPAEDEEVSVRPIGSAIYETFNTFNEDIWSIAHGWT-NG 320

Qy    231 NRVDLT--DKNIYSRDGMLILALTRKGQESFNGQVPRDDEPA 270
             : |  |  |   :|::  |: ::          ||| |
Db    321 QMFNATWYDSQITFSNGIMRFAIDKE----------DDEEA 351


RESULT 2
Q9APD8
ID   Q9APD8      PRELIMINARY;      PRT;   256 AA.
AC   Q9APD8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-beta-1,3-1,4-glucanase.
GN   BGLBC1.
OS   Bacillus circulans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1397;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC21367;
RA   Lee D.-S., Kim J.-Y., Kim H.-B.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF306531; AAG53947.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
SQ   SEQUENCE   256 AA;  27604 MW;  989E50C09F6E4531 CRC64;

  Query Match             14.0%;  Score 254.5;  DB 2;  Length 256;
  Best Local Similarity   34.6%;  Pred. No. 1.3e-09;
  Matches   62;  Conservative   29;  Mismatches   71;  Indels   17;  Gaps    7;

Qy     23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
          :| | | :: ||  | ::  ||: |||:| |  :| |:|:| |   |  | | |    |:||
Db     88 SNCSGKPYASAEYATTQKYGYGRVEARIKAAKGTGLVTSLFTY---SGAAPGTSNDEIDI 144

Qy     83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA--FHTYGLEWTPNYVRWTVDGQ 140
          |:|||:    ::|  |   |    |       :    | :  || |  ||:|  ::| |||:
Db    145 EILGKDTTKMETNYFTNGVGGHST-----VIDLGFDASLDFHDYAFEWSPTSIKWYVDGR 199

Qy    141 EVRKTEGGQVSNLTGTQG-LRFNLWSSESAA--WVGQFDESKLPLFQFINWVKVYKYTP 196
           |  || |    |  : | :  ||||    |  | |:|     |:  : :|:   |:||
Db    200 LVH-TETGSRGPLPTSPGYIMVNLWSGAGPAEIWTGKFTYPGHPIRAYYDWI---KFTP 254


RESULT 3
Q45648
ID   Q45648      PRELIMINARY;      PRT;   276 AA.
AC   Q45648;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Lichenase precursor.
GN   BGA1.
OS   Bacillus sp.
```

```
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1409;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=137;
RX   MEDLINE=94288605; PubMed=7517127;
RA   Tabernero C., Coll P.M., Fernandez-Abalos J.M., Perez P.,
RA   Santamaria R.I.;
RT   "Cloning and DNA sequencing of bgaA, a gene encoding an endo-beta-1,3-
RT   1, 4-glucanase, from an alkalophilic Bacillus strain (N137).";
RL   Appl. Environ. Microbiol. 60:1213-1220(1994).
DR   EMBL; Z12151; CAA78135.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Signal.
FT   SIGNAL        1     31       POTENTIAL.
FT   CHAIN        32    276       LICHENASE.
SQ   SEQUENCE   276 AA;  31770 MW;  D047F8A34CA9EBE2 CRC64;

  Query Match             13.6%;  Score 247.5;  DB 2;  Length 276;
  Best Local Similarity   36.1%;  Pred. No. 4.3e-09;
  Matches   61;  Conservative   20;  Mismatches   75;  Indels   13;  Gaps    6;

Qy     30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
          :   || | :   || ||  || | ::||||| | |  |    :  || |:||| |||:
Db     96 YKAGELRTNDFYHYGLFEVSMKPAKSTGTVSSFFTY-TGPWDWENDPWDEIDIEFLGKDT 154

Qy     90 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 149
             | |  |   |    :| :| :   |   |:||  || |  :|| |:|: |
Db    155 TKIQFNYFTNGVGG---NEHYHELGFDAADDFNTYAFEWRPESIRWFVNGELVHTA---- 207

Qy    150 VSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 196
           |:  | | :  |||       | |:|:    |:    :||   ||||
Db    208 TENIPQTPQKIMMNLWPGIGVDGWTGRFNGEDTPVVTQYDWV---KYTP 253


RESULT 4
Q93GE8
ID   Q93GE8      PRELIMINARY;      PRT;   214 AA.
AC   Q93GE8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254957; AAK50610.2; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   214 AA;  24159 MW;  3DB186D3B9D991CF CRC64;

  Query Match             13.5%;  Score 246.5;  DB 2;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 3.6e-09;
  Matches   61;  Conservative   25;  Mismatches   75;  Indels   11;  Gaps    6;

Qy     23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
          |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db     52 TSPSYNKFDCGENRSVQTYGYGPYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 108

Qy     83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
          | |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :
```

```
Db    109 EFLGKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 164

Qy    143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
          : |   |:    |   :  |||:      |:| ::    ||:   :||:  |
Db    165 KHTATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 213


RESULT 5
Q93GE7
ID   Q93GE7      PRELIMINARY;      PRT;   214 AA.
AC   Q93GE7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254959; AAK50612.2; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   214 AA;  24175 MW;  3DB192D2F6B4CAA3 CRC64;

  Query Match             13.5%;  Score 245.5;  DB 2;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 4.2e-09;
  Matches   61;  Conservative   25;  Mismatches   75;  Indels   11;  Gaps    6;

Qy     23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
          |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db     52 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 108

Qy     83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
          | |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :
Db    109 EFLGKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 164

Qy    143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
          : |   |:    |   :  |||:      |:| ::    ||:   :||:  |
Db    165 KHTATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 213


RESULT 6
Q9S310
ID   Q9S310      PRELIMINARY;      PRT;   802 AA.
AC   Q9S310;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Family 11 xylanase /family 16 beta (1,3-1,4) glucanase.
GN   XYND.
OS   Ruminococcus flavefaciens.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Lachnospiraceae; Ruminococcus.
OX   NCBI_TaxID=1265;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=17;
RA   Aurilia V., Martin J.C., Munro C.A., Mercer D.K., Flint H.J.;
RT   "Organisation and strain distribution of genes responsible for the
RT   utilization of xylans by the rumen cellulolytic bacterium Ruminococcus
RT   flavefaciens 17.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ132472; CAB51934.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR003305; CBM_CenC.
DR   InterPro; IPR001137; GH_11.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF02018; CBM_4_9; 2.
DR   Pfam; PF00457; Glyco_hydro_11; 2.
DR   Pfam; PF00722; Glyco_hydro_16; 2.
DR   PRINTS; PR00911; GLHYDRLASE11.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS00776; GLYCOSYL_HYDROL_F11_1; 2.
DR   PROSITE; PS00777; GLYCOSYL_HYDROL_F11_2; 2.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 2.
FT   CHAIN         2    802       FAMILY 11 XYLANASE /FAMILY 16 BETA
FT                                (1,3-1,4) GLUCANASE.
SQ   SEQUENCE   802 AA;  89019 MW;  F36BC68805FC5274 CRC64;

  Query Match             13.3%;  Score 242.5;  DB 2;  Length 802;
  Best Local Similarity   34.9%;  Pred. No. 3.7e-08;
  Matches   58;  Conservative   21;  Mismatches   74;  Indels   13;  Gaps    6;

Qy     30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
          :|| |  |     || :|  |:     | ||| | |   |   |  || |:|||:||||
Db    634 YSGGEFRTNNFYHYGYYECSMQAMKNDGVVSSFFTYTGPS---DDNPWDEIDIEILGKNT 690

Qy     90 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 149
              | |  |   |     || : :  : :|:|||| :| |||: | |||:|| :
Db    691 TQVQFNYYTNGQGKH---EKLYDLGFDSSEAYHTYGFDWQPNYIAWYVDGREVYRA---- 743

Qy    150 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
          ::  | | :  | |   :  |:  |:  : ||     ||   |
Db    744 TQDIPKTPGKIMMNAWPGLTVDDWLKAFN-GRTPLTAHYQWVTYNK 788


RESULT 7
Q93GE6
ID   Q93GE6      PRELIMINARY;      PRT;   214 AA.
AC   Q93GE6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254961; AAK50614.2; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   214 AA;  24131 MW;  1E9D8BFA4300EBA5 CRC64;

  Query Match             13.1%;  Score 239.5;  DB 2;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 1.1e-08;
  Matches   61;  Conservative   24;  Mismatches   76;  Indels   11;  Gaps    6;

Qy     23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
          |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db     52 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 108

Qy     83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
          | |||:    | |  |  ||     ||   :   |  |:| |  :| || ::| |||| :
Db    109 ESLGKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHMYAFDWQPNSIKWYVDGQ-L 164

Qy    143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
          : |   ||    |   :  |||:      |:| ::    ||:   :||:  |
Db    165 KHTATTQVPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 213


RESULT 8
Q45691
ID   Q45691      PRELIMINARY;      PRT;   239 AA.
```

```
AC   Q45691;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-beta-1,3-1,4-glucanase.
GN   BEG1.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCIB 8565;
RA   van Rensburg P., van Zyl W.H., Pretorius I.S.;
RT   "Over-expression of the Saccharomyces cerevisiae exo-beta-1,3-
RT   glucanase gene together with the Butyrivibrio fibrisolvens endo-beta-
RT   1,4-glucanase gene and the Bacillus subtilis endo-beta-1,3-1,4-
RT   glucanase gene in Saccharomyces cerevisiae.";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U60830; AAB05759.1; -.
DR   HSSP; P27051; 1GBG.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
SQ   SEQUENCE   239 AA;  26969 MW;  C4219760D13F878A CRC64;

  Query Match             13.1%;  Score 238.5;  DB 2;  Length 239;
  Best Local Similarity   33.7%;  Pred. No. 1.4e-08;
  Matches   59;  Conservative   27;  Mismatches   72;  Indels   17;  Gaps    7;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db      77 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 133

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPA---ADQAFHTYGLEWTPNYVRWTVDG 139
           | |||:    | |  |  |      | |  ::     |  |:|||  :| || ::| ||
Db     134 EFLGKDTTKVQFNYYTNGA------ENHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDW 187

Qy     140 QEVRKTEGGQVSNLTGTQGLRFNLWSSESAA-WVGQFDESKLPLFQFINWVKVYK 193
           | :: |   |:    |   :  |||:    : |:| ::    ||:   :||:  |
Db     188 Q-LKHTATTQIPTAPGK--IMMNLWNGTGVSDWLGSYNGVN-PLYAHYDWVRYTK 238


RESULT 9
Q8RMP0
ID   Q8RMP0      PRELIMINARY;      PRT;   242 AA.
AC   Q8RMP0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Beta-1,3-1,4-glucanase.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sun J., Li W., Gu S., Xu Z., Zhao H., Xiao J.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF490978; AAM08358.1; -.
SQ   SEQUENCE   242 AA;  27238 MW;  285FED3FE76AE69A CRC64;

  Query Match             13.0%;  Score 236.5;  DB 2;  Length 242;
  Best Local Similarity   34.9%;  Pred. No. 2e-08;
  Matches   60;  Conservative   25;  Mismatches   76;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||
Db      80 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDI 136

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :
Db     137 EFLGKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 192

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           : |   |:    |   :  |||:       :| ::    ||:   :||:  |
Db     193 KHTATNQIPTTPGK--IMMNLWNGTGVDERLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 10
Q9FDC9
ID   Q9FDC9      PRELIMINARY;      PRT;   212 AA.
AC   Q9FDC9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-1,3-1,4-beta-glucanase (Fragment).
OS   Paenibacillus polymyxa (Bacillus polymyxa).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Paenibacillaceae; Paenibacillus.
OX   NCBI_TaxID=1406;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yao W., Wang Y., Song W., Yang K., Su Z.;
RT   "Gene cloning of an antifugal protein.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF284449; AAG02415.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
FT   NON_TER       1      1
FT   NON_TER     212    212
SQ   SEQUENCE   212 AA;  24127 MW;  7E707E7FBEF7A440 CRC64;

  Query Match             12.7%;  Score 232.5;  DB 2;  Length 212;
  Best Local Similarity   34.8%;  Pred. No. 3.1e-08;
  Matches   62;  Conservative   26;  Mismatches   73;  Indels   17;  Gaps    9;

Qy      20 ALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVE 79
           :||::   | | | |  :    :|| :|  || |  :| ||| | |      |:|  | |
Db      48 SLTSSAYNK-FDGGEYRSKNTYRYGLYEVNMKPAKNTGIVSSFFTYTGP---ANGTQWDE 103

Qy      80 VDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 139
           :||| |||:    | |  |  |     ||   :   |  ||||  :| | |::| |||
Db     104 IDIEFLGKDTTKVQFNYYTNGIGGH---EKVVDLGFDASSGFHTYAFDWQPGYIKWYVDG 160

Qy     140 QEVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           :: |     :|:  | | :  |||:     :|:| ::    ||:   :||   |||
Db     161 -VLKHT---ATTNIPKTPGQIMMNLWNGTGVDSWLGPYNGVN-PLYAEYDWV---KYT 210


RESULT 11
O07856
ID   O07856      PRELIMINARY;      PRT;   237 AA.
AC   O07856;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Beta-(1,3-1,4)-glucanase precursor.
GN   LICHENASE.
OS   Streptococcus bovis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Streptococcus.
OX   NCBI_TaxID=1315;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JB1;
RA   Ekinci M., Flint H.J.;
RT   "Isolation and overexpression of a gene encoding an extracellular
RT   beta-(1,3-1,4)-glucanase from Streptococcus bovis JB1.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z92911; CAB07443.1; -.
DR   HSSP; P23904; 1AJK.
```

```
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    237       BETA-(1,3-1,4)-GLUCANASE.
SQ   SEQUENCE   237 AA;  26989 MW;  7DEF5BCE53790470 CRC64;

  Query Match             12.6%;  Score 230;  DB 2;  Length 237;
  Best Local Similarity   33.1%;  Pred. No. 5.2e-08;
  Matches   54;  Conservative   26;  Mismatches   69;  Indels   14;  Gaps    6;

Qy      30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
           ::| |  : |   || |:  ||     | ||| | |   |   ||  | |:||| |||:
Db      81 YTGGEWRSKERFGYGLFQVNMKPIKNPGVVSSFFTYTGPS---DGTKWDEIDIEFLGKDT 137

Qy      90 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 149
             | |  |     |   | : :   | | ||||| :|  ::: | |||: |
Db     138 TKVQFNYYT---SGQGNHEYLYNLGFDASQGFHTYGFDWQADHITWYVDGRAVYTA---- 190

Qy     150 VSNLTGTQG-LRFNLW--SSESAAWVGQFDESKLPLFQFINWV 189
           :|:  | | :  | |  : |  :|:| ::  : ||: : :|:
Db     191 YNNIPSTPGKIMMNAWPGTHEVDSWLGAYN-GRTPLYAYYDWI 232


RESULT 12
Q93UI1
ID   Q93UI1      PRELIMINARY;      PRT;   205 AA.
AC   Q93UI1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254960; AAK50613.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER     205    205
SQ   SEQUENCE   205 AA;  22890 MW;  E7CADAAFE02A2F51 CRC64;

  Query Match             12.2%;  Score 223;  DB 2;  Length 205;
  Best Local Similarity   36.1%;  Pred. No. 1.3e-07;
  Matches   56;  Conservative   21;  Mismatches   68;  Indels   10;  Gaps    5;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db      52 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 108

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  | ||     ||   :   |  |:| |  :| || ::| |||| :
Db     109 ESLGKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHMYAFDWQPNSIKWYVDGQ-L 164

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 176
           : |   ||    |   :  |||:       |:| ::
Db     165 KHTATTQVPAAPGK--IMMNLWNGTGVDDWLGSYN 197


RESULT 13
Q97FD3
ID   Q97FD3      PRELIMINARY;      PRT;   246 AA.
AC   Q97FD3;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-1,3(4)-beta-glucanase family 16.
GN   CAC2807.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
DR   EMBL; AE007778; AAK80751.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Complete proteome.
SQ   SEQUENCE   246 AA;  27717 MW;  C0EBC302678D2FE1 CRC64;

  Query Match             12.1%;  Score 221;  DB 16;  Length 246;
  Best Local Similarity   33.7%;  Pred. No. 2.2e-07;
  Matches   58;  Conservative   19;  Mismatches   73;  Indels   22;  Gaps    8;

Qy      30 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 89
           ::| |  :     || :   || |   |  || | |   |   |  || |:||| |||:
Db      91 YAGGEYRSNNRYGYGLYRVSMKPAKHIGVDSSFFSYTGPS---DNNPWDEIDIEFLGKDT 147

Qy      90 GSFQSNIITGKAGAQKTSEKH---HAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 146
             | |  |   |       ||   : :   | : ||||  |  ||: | |||:|| :
Db     148 TEVQFNYYTNGVG------KHEYLYKLGFDASKGFHTYGYIWEQNYIAWLVDGKEVYRA- 200

Qy     147 GGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 196
             ||:    | :  |||      :|:| :|    |:  : ||     | |
Db     201 ---TSNIPTHPGKVMMNLWPGIGVDSWLGAYD-GVTPVKAYYNWA---MYNP 245


RESULT 14
Q8U8N5
ID   Q8U8N5      PRELIMINARY;      PRT;   289 AA.
AC   Q8U8N5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Endo-1,3-1,4-beta-glycanase.
GN   EXOK OR ATU4055 OR AGR_L_1600.
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193;
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.;
RT   "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT   C58.";
RL   Science 294:2317-2323(2001).
RN   [2]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608551; PubMed=11743194;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
DR   EMBL; AE009336; AAL44856.1; ALT_INIT.
DR   EMBL; AE008279; AAK89373.1; -.
KW   Complete proteome.
SQ   SEQUENCE   289 AA;  32368 MW;  0384C6F83320EAC9 CRC64;

  Query Match             11.3%;  Score 206;  DB 16;  Length 289;
  Best Local Similarity   35.1%;  Pred. No. 2.7e-06;
  Matches   54;  Conservative   26;  Mismatches   56;  Indels   18;  Gaps    6;

Qy      28 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 87
           ::|:  |: |    :|| :||||| |  ||  |: | |   :   | :|  |:| |||||
Db     115 RNFACGEIQTKGRYRYGTYEARMKAATGSGLNSAFFTYIGPT---DKKPHDEIDFEVLGK 171

Qy      88 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 146
           | |  | |  |  | |    :||   |   ||  |: |   | |  :|: |:|: |
Db     172 NTGKVQLNQYIAAKGG----NEKLVPVEGGADAGFNDYAFVWEPQRLRYYVNGKLVH--- 224

Qy     147 GGQVSNLT----GTQGLRFNLWSSESAA-WVGQF 175
           :|:: |      | : |:|| :::   |:|:|
Db     225 --EVTDETKIPQNAQKIFFSLWGTDTLKDWMGKF 256


RESULT 15
Q98C78
ID   Q98C78      PRELIMINARY;      PRT;   293 AA.
AC   Q98C78;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Endo-1,3-1,4-beta-glycanase, ExoK.
GN   MLR5264.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003006; BAB51743.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
KW   Complete proteome.
SQ   SEQUENCE   293 AA;  32347 MW;  93BBA1672176ED3C CRC64;

  Query Match             11.0%;  Score 200;  DB 16;  Length 293;
  Best Local Similarity   33.8%;  Pred. No. 7e-06;
  Matches   51;  Conservative   27;  Mismatches   61;  Indels   12;  Gaps    6;

Qy      28 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 87
           ::|:  |: | :   || :|||:|    ||  :: | |   |   | :|| |:| |:| |
Db     112 REFACGEIQTKQRFGYGTYEARLKTDTGSGLNAAFFTYIGPS---DKQPWDEIDFEILTK 168

Qy      88 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK-T 145
           :   | |  | ||   :|  |    |    |:||: |   |  : :|| |:|| |   |
Db     169 DTSKVQVNAYIDGKGKNEKLVE----VPGGTDKAFNDYAFVWEKDSLRWYVNGQLVNTIT 224

Qy     146 EGGQVSNLTGTQGLRFNLWSSES-AAWVGQF 175
           :  ::   :  | : |:|| ||:   |:| |
Db     225 DPAKLP--SHAQKIFFSLWGSETMKGWMGAF 253


Search completed: January  9, 2003, 12:18:13
Job time : 32.103 secs
```

```
DR   PDB; 1CPN; 22-JUN-94.
DR   PDB; 1MAC; 27-FEB-95.
DR   PDB; 1AJK; 06-MAY-98.
DR   PDB; 1AJO; 06-MAY-98.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Hydrolase; Glycosidase; Signal; 3D-structure.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    237       BETA-GLUCANASE.
FT   ACT_SITE    128    128       NUCLEOPHILE (BY SIMILARITY).
FT   ACT_SITE    132    132       PROTON DONOR (BY SIMILARITY).
FT   DISULFID     55     84
FT   MUTAGEN     128    128       E->D,N,A,L,P,R,H,C,S,Y: LOSS OF ACTIVITY.
FT   STRAND       26     26
FT   STRAND       29     31
FT   STRAND       41     43
FT   TURN         51     52
FT   STRAND       53     53
FT   STRAND       56     57
FT   HELIX        59     61
FT   STRAND       62     64
FT   TURN         66     67
FT   STRAND       70     78
FT   TURN         79     80
FT   STRAND       81     89
FT   STRAND       93     93
FT   STRAND       96    103
FT   TURN        108    109
FT   STRAND      110    118
FT   HELIX       120    122
FT   TURN        123    123
FT   STRAND      127    134
FT   TURN        135    136
FT   STRAND      137    147
FT   TURN        148    149
FT   STRAND      150    150
FT   STRAND      155    158
FT   TURN        163    164
FT   STRAND      168    174
FT   STRAND      179    183
FT   TURN        184    185
FT   STRAND      186    191
FT   STRAND      200    209
FT   TURN        213    216
FT   STRAND      225    236
SQ   SEQUENCE   237 AA;  26589 MW;  436EABCDFFC87781 CRC64;

  Query Match          12.5%;  Score 228;  DB 1;  Length 237;
  Best Local Similarity   34.3%;  Pred. No. 1.5e-09;
  Matches   60;  Conservative   25;  Mismatches   74;  Indels   16;  Gaps    8;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: :   |  ||  :     || :|  || |  :| ||| | |      | |  | |:||
Db      75 TSSAYNKFDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDI 131

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:   .| |  |   |     ||  ::   | : ||||  :| | |::| |||  :
Db     132 EFLGKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VL 187

Qy     143 RKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           : |     :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     188 KHT---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 235


RESULT 10
GUB_PAEPO
ID   GUB_PAEPO      STANDARD;      PRT;   238 AA.
AC   P45797;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE   (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN   GLUB.
OS   Paenibacillus polymyxa (Bacillus polymyxa).
OC   Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
OX   NCBI_TaxID=1406;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 842;
RX   MEDLINE=92041687; PubMed=1938968;
RA   Gosalbes M.J., Perez-Gonzalez J.A., Gonzalez R., Navarro A.;
RT   "Two beta-glycanase genes are clustered in Bacillus polymyxa:
RT   molecular cloning, expression, and sequence analysis of genes
RT   encoding a xylanase and an endo-beta-(1,3)-(1,4)-glucanase.";
RL   J. Bacteriol. 173:7705-7710(1991).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC       in beta-D-glucans containing 1,3- and 1,4-bonds.
CC   -!- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC       SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X57094; CAA40379.1; -.
DR   HSSP; P23904; 1CPN.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Hydrolase; Glycosidase; Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    238       BETA-GLUCANASE.
FT   ACT_SITE    129    129       NUCLEOPHILE (BY SIMILARITY).
FT   ACT_SITE    133    133       PROTON DONOR (BY SIMILARITY).
FT   DISULFID     56     85       BY SIMILARITY.
SQ   SEQUENCE   238 AA;  26919 MW;  C0CF7B4EA5D40E8C CRC64;

  Query Match          12.5%;  Score 227.5;  DB 1;  Length 238;
  Best Local Similarity   34.3%;  Pred. No. 1.6e-09;
  Matches   61;  Conservative   26;  Mismatches   74;  Indels   17;  Gaps    9;

Qy      20 ALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVE 79
           :||:  : | |   |  :     || :|  || |  :| ||| | |   |    | | |
Db      74 SLTSPANNK-FDCGEYRSTNNYGYGLYEVSMKPAKNTGIVSSFFTYTGPSH---GTQWDE 129

Qy      80 VDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 139
           :||| |||:    | |  |   |     ||   :   |  :|||| :| | |::| |||
Db     130 IDIEFLGKDTTKVQFNYYTNGVGGH---EKIINLGFDASTSFHTYAFDWQPGYIKWYVDG 186

Qy     140 QEVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           :: |     :|:  | | :  |||:      :|:| :: :  ||:   :||   |||
Db     187 -VLKHT---ATTNIPSTPGKIMMNLWNGTGVDSWLGSYNGAN-PLYAEYDWV---KYT 236


RESULT 11
EXOK_RHIME
ID   EXOK_RHIME     STANDARD;      PRT;   269 AA.
AC   P33693;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Endo-1,3-1,4-beta-glycanase exoK precursor (EC 3.2.1.-) (Succinoglycan
DE   biosynthesis protein exoK).
GN   EXOK OR RB1080 OR SMB20955.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OG   Plasmid pSymB (megaplasmid 2).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae grou
```

```
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=94042869; PubMed=8226645;
RA   Glucksmann M.A., Reuber T.L., Walker G.C.;
RT   "Family of glycosyl transferases needed for the synthesis of
RT   succinoglycan by Rhizobium meliloti.";
RL   J. Bacteriol. 175:7033-7044(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RCR2011 / SU47;
RX   MEDLINE=93241147; PubMed=8479421;
RA   Becker A., Kleickmann A., Arnold W., Puehler A.;
RT   "Analysis of the Rhizobium meliloti exoH/exoK/exoL fragment: ExoK
RT   shows homology to excreted endo-beta-1,3-1,4-glucanases and ExoH
RT   resembles membrane proteins.";
RL   Mol. Gen. Genet. 238:145-154(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396508; PubMed=11481431;
RA   Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA   Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA   Golding B., Puehler A.;
RT   "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT   fixing endosymbiont Sinorhizobium meliloti.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=98226741; PubMed=9560202;
RA   York G.M., Walker G.C.;
RT   "The Rhizobium meliloti ExoK and ExsH glycanases specifically
RT   depolymerize nascent succinoglycan chains.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:4912-4917(1998).
CC   -!- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
CC       SUCCINOGLYCAN. DYNAMICALLY REGULATE THE MOLECULAR WEIGHT
CC       DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
CC       ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS, PERHAPS BEFORE
CC       IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
CC       AGGREGATION STATE.
CC   -!- PATHWAY: Exopolysaccharide biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L20758; AAA16048.1; -.
DR   EMBL; Z17219; CAA78927.1; -.
DR   EMBL; AL603645; CAC49480.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Exopolysaccharide synthesis; Glycosidase; Hydrolase; Signal; Plasmid;
KW   Complete proteome.
FT   SIGNAL       1     27       POTENTIAL.
FT   CHAIN       28    269       ENDO-1,3-1,4-BETA-GLYCANASE EXOK.
FT   CONFLICT    93    102       ACGEIQTRKR -> LRRNPDAQG (IN REF. 2).
SQ   SEQUENCE   269 AA;  30083 MW;  073CC7ED65EF2611 CRC64;

  Query Match             10.8%;  Score 197;  DB 1;  Length 269;
  Best Local Similarity   32.5%;  Pred. No. 2.8e-07;
  Matches   50;  Conservative   31;  Mismatches   61;  Indels   12;  Gaps    6;

Qy      25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84
           |  ::|:  |: | :   || :|||:| |  ||  |: | |      || :|  |:| ||
Db      87 VKERNFACGEIQTRKRFGYGTYEARIKAADGSGLNSAFFTYIGP---ADKKPHDEIDFEV 143

Qy      85 LGKNPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 143
           ||||    | |  :: | | :  ::    |   |:| |: |   |  | :|: |:|: |
Db     144 LGKNTAKVQINQYVSAKGGNEFLAD----VPGGANQGFNDYAFVWEKNRIRYYVNGELVH 199

Qy     144 K-TEGGQVSNLTGTQGLRFNLWSSESAA-WVGQF 175
           : |:  ::      | : |:|| :::   |:| |
Db     200 EVTDPAKIP--VNAQKIFFSLWGTDTLTDWMGTF 231


RESULT 12
UTR2_YEAST
ID   UTR2_YEAST     STANDARD;      PRT;   347 AA.
AC   P32623;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   UTR2 protein (Unknown transcript 2 protein).
GN   UTR2 OR YEL040W OR SYGP-ORF18.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
RA   Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RA   Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA   Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA   Chung E., Duncan M., Guzman E., Hartzell G., Hunicke-Smith S.,
RA   Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA   Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA   Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA   Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B-6441;
RX   MEDLINE=94016558; PubMed=8411151;
RA   Melnick L., Sherman F.;
RT   "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
RT   of Saccharomyces cerevisiae share a common ancestry.";
RL   J. Mol. Biol. 233:372-388(1993).
CC   -!- SIMILARITY: SOME, TO YEAST YGR189C.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U18779; AAB65002.1; ALT_INIT.
DR   EMBL; L22173; AAA34941.1; -.
DR   EMBL; S65964; AAD13975.1; -.
DR   EMBL; S66130; AAB28444.1; -.
DR   PIR; S30839; S30839.
DR   HSSP; P23904; 1AJO.
DR   SGD; S0000766; UTR2.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
FT   DOMAIN      234    322       SER-RICH.
FT   DOMAIN      269    283       POLY-SER.
FT   CONFLICT     10     10       L -> V (IN REF. 3).
FT   CONFLICT    171    171       A -> R (IN REF. 3).
```

```
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M63166; AAA32828.1; ALT_FRAME.
DR   EMBL; D63508; BAA09783.1; -.
DR   EMBL; AL109796; CAB52471.1; -.
DR   EMBL; AL161576; CAB81020.1; -.
DR   EMBL; AY035156; AAK59660.1; -.
DR   EMBL; AY063027; AAL34201.1; -.
DR   EMBL; Z17602; CAA79012.1; ALT_FRAME.
DR   EMBL; AF035384; AAC39467.1; -.
DR   PIR; JQ1022; JQ1022.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Glycosidase; Hydrolase; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    269       MERI-5 PROTEIN.
FT   CONFLICT     93     94       GS -> DR (IN REF. 7).
FT   CONFLICT    184    184       A -> G (IN REF. 1).
SQ   SEQUENCE   269 AA;  30755 MW;  648F042BC7ADED86 CRC64;

  Query Match             8.6%;  Score 157.5;  DB 1;  Length 269;
  Best Local Similarity   25.2%;  Pred. No. 0.00019;
  Matches   56;  Conservative   39;  Mismatches   76;  Indels   51;  Gaps   10;

Qy       6 TAVKSALAVAAAAAALTTNVS-----AKDFSGAELYTLE-------------EVQYGKFE 47
           | : :| :|:||      ||:       |  :  :| ||             |  :|| :
Db      11 TLLVAAFSVSAADFNTDVNVAWGNGRGKILNNGQLLTLSLDKSSGSGFQSKTEYLFGKID 70

Qy      48 ARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK---NPGSFQSNIITGKAG 102
             ::|:    ::|||:: :|   ||       | |:| | ||    :| :  :|: |   |
Db      71 MQIKLVPGNSAGTVTTFYLKSEGS------TWDEIDFEFLGNMSGDPYTLHTNVYTQGKG 124

Qy     103 AQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRF- 161
           |  : |    | |:  ||| : | |  :  |||   :|: :  :      : |: |
Db     125 -DKEQQFHLWFDPTAN--FHTYSILWNPQRIILTVDDTPIREFKNYE------SLGVLFP 175

Qy     162 ---------NLWSSESAAWVG---QFDESKLPLFQFINWVKV 191
                   :||:::  |  |   : | || |       :|:
Db     176 KNKPMRMYASLWNADDWATRGGLVKTDWSKAPFMASYRNIKI 217


RESULT 14
GUB_RHOMR
ID   GUB_RHOMR      STANDARD;      PRT;   286 AA.
AC   P45798;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE   (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN   BGLA.
OS   Rhodothermus marinus (Rhodothermus obamensis).
OC   Bacteria; Bacteroidetes; Sphingobacteria; Sphingobacteriales;
OC   Crenotrichaceae; Rhodothermus.
OX   NCBI_TaxID=29549;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=21 / ITI-378;
RX   MEDLINE=95010084; PubMed=7925416;
RA   Spilliaert R., Hreggvidsson G.O., Kristjansson J.K.,
RA   Eggertsson G., Palsdottir A.;
RT   "Cloning and sequencing of a Rhodothermus marinus gene, bglA, coding
RT   for a thermostable beta-glucanase and its expression in Escherichia
RT   coli.";
RL   Eur. J. Biochem. 224:923-930(1994).
CC   -!- FUNCTION: SHOWS ACTIVITY ON LICHENAN, BETA-GLUCAN AND LAMINARIN
CC       BUT NOT ON CMC CELLULOSE OR XYLAN. THE ENZYME HAS A TEMPERATURE
CC       OPTIMUM OF 85 DEGREES CELSIUS AND A PH OPTIMUM OF 7.0.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC       in beta-D-glucans containing 1,3- and 1,4-bonds.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U04836; AAA60459.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Hydrolase; Glycosidase; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    286       BETA-GLUCANASE.
FT   ACT_SITE    158    158       NUCLEOPHILE (BY SIMILARITY).
FT   ACT_SITE    163    163       PROTON DONOR (BY SIMILARITY).
SQ   SEQUENCE   286 AA;  33145 MW;  7215C33624135191 CRC64;

  Query Match             7.3%;  Score 133.5;  DB 1;  Length 286;
  Best Local Similarity   23.7%;  Pred. No. 0.01;
  Matches   44;  Conservative   34;  Mismatches   85;  Indels   23;  Gaps    8;

Qy      28 KDFSGAELYTLEEVQ--YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWV---EVDI 82
           :::: | | |  :    ||:|| | :: :  ||  ::::  :          |    |:||
Db     103 REYTSARLVTRGKASWTYGRFEIRARLPSGRGTWPAIWMLPDRQTYGSAY-WPDNGEIDI 161

Qy      83 -EVLGKNP----GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTV 137
            | :| ||    |:  :       | |:         | |   || | :||||  :|| |
Db     162 MEHVGFNPDVVHGTVHTKAYNHLLGTQRGGSIR---VPTARTDFHVYAIEWTPEEIRWFV 218

Qy     138 DGQEVRKTEGGQVSNLTG-------TQGLRFNLWSSESAAWVGQ--FDESKLPLFQFINW 188
           |      :   ::::          |     :  :   || ||   |    |     :::
Db     219 DDSLYYRFPNERLTDPEADWRHWPFDQPFHLIMNIAVGGAWGGQQGVDPEAFPAQLVVDY 278

Qy     189 VKVYKY 194
           |:||::
Db     279 VRVYRW 284


RESULT 15
BRU1_SOYBN
ID   BRU1_SOYBN     STANDARD;      PRT;   283 AA.
AC   P35694;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Brassinosteroid-regulated protein BRU1 precursor.
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Epicotyl;
RX   MEDLINE=94159788; PubMed=8115544;
RA   Zurek D.M., Clouse S.D.;
RT   "Molecular cloning and characterization of a brassinosteroid-regulated
RT   gene from elongating soybean (Glycine max L.) epicotyls.";
RL   Plant Physiol. 104:161-170(1994).
CC   -!- FUNCTION: POSSIBLE ROLE IN BRASSINOSTEROID-STIMULATED ELONGATION.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
FT   CONFLICT    234    234       S -> C (IN REF. 3).
SQ   SEQUENCE   347 AA;  36692 MW;  1E8AFB862C4BB328 CRC64;

  Query Match             10.4%;  Score 189.5;  DB 1;  Length 347;
  Best Local Similarity   24.6%;  Pred. No. 1.3e-06;
  Matches   83;  Conservative   49;  Mismatches  120;  Indels   85;  Gaps   15;

Qy      28 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 87
           |:  |  | :   | |||  ||:| :  :| |:   ||    :         |:| | :|
Db       3 KNSGGTVLSSTRAVWYGKVSARIKTSHLAGVVTGFILYSGAGD--------ELDYEFVGA 54

Qy      88 NPGSFQSNI-------ITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQ 140
           :  : |:|         |  |    |          : :||| |:|  :|| |::||
Db      55 DLETAQTNFYWESVLNYTNSANISTTD---------TFENYHTYELDWHEDYVTWSIDGV 105

Qy     141 EVRKTEGGQVSNLTGTQGLRF---------NLW---SSESA----AWVG---QFDESKLP 181
             |     :  | | ||  ::          ::|   :| :|    || |    :| | :
Db     106 VGRTLYKNETYNAT-TQKYQYPQTPSKVDISIWPGGNSTNAPGTIAWSGGEINWDASDIS 164

Qy     182 ----LFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTD 237
                :  :| | :  | |      ||   : |  :     |        |    : :||
Db     165 NPGYYYAIVNEVNITCYDP-----PSDTKKNGTSAYVYTSSSE-------FLAKDIAITD 212

Qy     238 KNIY--SRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPAS-------- 287
             :   | :|        | :   |       : | |||: |:|||  :|
Db     213 DEVMMDSDEG--------SGLDPHKGAT----TSSTQKSSSSTATSSSKTSSDHSSSTKK 260

Qy     288 ---SSSVPASSSSAFVPPSSSSATNAIHGMRTTPAVA 321
              :||  :||||:    ||||:|   :| :   :|:
Db     261 SSKTSSTASSSSSSSSSSSSSSTATKNGDKVVSSVS 297


RESULT 13
MER5_ARATH
ID   MER5_ARATH     STANDARD;      PRT;   269 AA.
AC   P24806; Q39148; Q41904; O64956;
DT   01-MAR-1992 (Rel. 21, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   MERI-5 protein precursor (Endo-xyloglucan transferase) (Xyloglucan
DE   endo-1,4-beta-D-glucanase).
GN   MERI-5 OR MERI5B OR SEN4 OR AT4G30270 OR F9N11.120.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93005704; PubMed=1840916;
RA   Medford J.I., Elmer J.S., Klee H.J.;
RT   "Molecular cloning and characterization of genes expressed in shoot
RT   apical meristems.";
RL   Plant Cell 3:359-370(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Kamimai T., Tomita E., Nishitani K.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Arrowsmith D.A.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT   SSP consortium (Salk/Stanford/PGEC).";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 1-132 FROM N.A.
RC   STRAIN=cv. Columbia; TISSUE=Green siliques;
RA   Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.;
RL   Submitted (OCT-1992) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 1-120 FROM N.A.
RC   STRAIN=cv. Columbia; TISSUE=Leaf;
RX   MEDLINE=98278374; PubMed=9617812;
RA   Park J.-H., Oh S.A., Kim Y.H., Woo H.R., Nam H.G.;
RT   "Differential expression of senescence-associated mRNAs during leaf
RT   senescence induced by different senescence-inducing factors in
RT   Arabidopsis.";
RL   Plant Mol. Biol. 37:445-454(1998).
CC   -!- FUNCTION: Involved in cell wall reconstruction.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN SHOOT APICAL MERISTEMS, ALSO
CC       FOUND IN SEEDLINGS AND MERISTEMS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in positions 158; 178; 183; 189; 190; 194 and 199.
CC   -!- CAUTION: Ref.6 sequence differs from that shown due to frameshifts
CC       in positions 93 and 104.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L22162; AAA81350.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Glycosidase; Hydrolase; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    283       BRASSINOSTEROID-REGULATED PROTEIN BRU1.
SQ   SEQUENCE   283 AA;  32254 MW;  C248810EC7835737 CRC64;

  Query Match             7.0%;  Score 128;  DB 1;  Length 283;
  Best Local Similarity   23.7%;  Pred. No. 0.026;
  Matches   46;  Conservative   41;  Mismatches   73;  Indels   34;  Gaps    9;

Qy      27 AKDFSGAELYTL-------------EEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEI 71
           || |:| :| :|              :|  :|: : ::|: |  ::|||:: :|
Db      45 AKIFNGGQLLSLSLDKVSGSGFKSKKEYLFGRIDMQLKLVAGNSAGTVTAYYL------S 98

Qy      72 ADGRPWVEVDIEVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEW 128
           : |    |:| | ||    :|    :|| |   |  :  : :    |  :  |||| : |
Db      99 SQGPTHDEIDFEFLGNLSGDPYILHTNIFTQGKG-NREQQFYLWFDPTRN--FHTYSIIW 155

Qy     129 TPNYVRWTVDGQEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLP 181
            | :: : ||   :|  :  :   :     | :|   :||:::  |  |   : | || |
Db     156 KPQHIIFLVDNTPIRVFKNAEPLGVPFPKNQPMRIYSSLWNADDWATRGGLVKTDWSKAP 215

Qy     182 LFQFINWVKVYKYT 195
           :    |  :::
Db     216 FTAYYRNFKAIEFS 229


Search completed: January  9, 2003, 12:16:49
Job time : 9.88657 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:50:42 ; Search time 115.824 Seconds
(without alignments)
1368.672 Million cell updates/sec

Title: US-09-654-652A-1
Perfect score: 1333
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....TRKGQESFNGQVPRDDEPAP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1333	100.0	248	20	US-09-654-652A-1	Sequence 1, Appli
2	1333	100.0	267	20	US-09-654-652A-2	Sequence 2, Appli
3	1328	99.6	349	20	US-09-654-652A-3	Sequence 3, Appli
4	1328	99.6	349	21	US-09-791-537-30441	Sequence 30441, A
5	272.5	20.4	214	21	US-09-791-537-145019	Sequence 145019,
6	266.5	20.0	259	21	US-09-791-537-106873	Sequence 106873,
7	265.5	19.9	252	21	US-09-791-537-32335	Sequence 32335, A
8	257.5	19.3	308	18	US-09-463-862A-1	Sequence 1, Appli
9	256.5	19.2	208	21	US-09-791-537-6910	Sequence 6910, Ap
10	250.5	18.8	394	21	US-09-791-537-132244	Sequence 132244,
11	247.5	18.6	214	21	US-09-791-537-117361	Sequence 117361,
12	247.5	18.6	276	21	US-09-791-537-31771	Sequence 31771, A
13	246.5	18.5	242	21	US-09-791-537-27063	Sequence 27063, A
14	245.5	18.4	242	21	US-09-791-537-73450	Sequence 73450, A
15	245.5	18.4	242	21	US-09-791-537-79889	Sequence 79889, A
16	245	18.4	239	21	US-09-791-537-103347	Sequence 103347,
17	244	18.3	239	21	US-09-791-537-103324	Sequence 103324,
18	242.5	18.2	802	21	US-09-791-537-50107	Sequence 50107, A
19	242.5	18.2	802	21	US-09-791-537-116925	Sequence 116925,
20	238.5	17.9	239	21	US-09-791-537-30436	Sequence 30436, A
21	237	17.8	334	21	US-09-791-537-48909	Sequence 48909, A
22	235.5	17.7	214	21	US-09-791-537-88552	Sequence 88552, A
23	234.5	17.6	239	21	US-09-791-537-18679	Sequence 18679, A
24	233	17.5	239	21	US-09-791-537-103817	Sequence 103817,
25	231.5	17.4	320	21	US-09-791-537-67048	Sequence 67048, A
26	230	17.3	237	21	US-09-791-537-72848	Sequence 72848, A
27	229.5	17.2	243	21	US-09-791-537-30440	Sequence 30440, A
28	226	17.0	212	21	US-09-791-537-88873	Sequence 88873, A
29	226	17.0	214	21	US-09-791-537-76270	Sequence 76270, A
30	226	17.0	237	21	US-09-791-537-103320	Sequence 103320,
31	226	17.0	237	21	US-09-791-537-106876	Sequence 106876,
32	226	17.0	239	21	US-09-791-537-103345	Sequence 103345,
33	226	17.0	239	21	US-09-791-537-103350	Sequence 103350,
34	226	17.0	239	21	US-09-791-537-103355	Sequence 103355,
35	226	17.0	239	21	US-09-791-537-103357	Sequence 103357,
36	226	17.0	239	21	US-09-791-537-103814	Sequence 103814,
37	226	17.0	239	21	US-09-791-537-103819	Sequence 103819,
38	226	17.0	240	21	US-09-791-537-64827	Sequence 64827, A
39	225	16.9	238	21	US-09-791-537-79887	Sequence 79887, A
40	222.5	16.7	242	21	US-09-791-537-37095	Sequence 37095, A
41	222	16.7	237	21	US-09-791-537-3485	Sequence 3485, Ap
42	214	16.1	160	21	US-09-706-729-14	Sequence 14, Appl
43	206	15.5	220	21	US-09-739-449-8386	Sequence 8386, Ap
44	206	15.5	220	22	US-09-803-110-8386	Sequence 8386, Ap
45	204.5	15.3	504	27	US-60-385-568-329	Sequence 329, App

ALIGNMENTS

RESULT 1
US-09-654-652A-1
; Sequence 1, Application US/09654652A
; GENERAL INFORMATION:
; APPLICANT: SHYUR, LIE-FEN
; APPLICANT: CHEN, JUI-LIN
; APPLICANT: YANG, NING-SUN
; TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1,
; TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
; TITLE OF INVENTION: THERMO-TOLERANCE
; FILE REFERENCE: 4910-8
; CURRENT APPLICATION NUMBER: US/09/654,652A
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme
; OTHER INFORMATION: with enhanced activity and thermal stability
US-09-654-652A-1

Query Match 100.0%; Score 1333; DB 20; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.4e-130;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60

Qy     61 VLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120

Qy    121 KTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 KTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS 180

Qy    181 DFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQV 240

Qy    241 PRDDEPAP 248
          ||||||||
Db    241 PRDDEPAP 248


RESULT 2
US-09-654-652A-2
; Sequence 2, Application US/09654652A
; GENERAL INFORMATION:
;  APPLICANT: SHYUR, LIE-FEN
;  APPLICANT:  CHEN, JUI-LIN
;  APPLICANT:  YANG, NING-SUN
;  TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1,
;  TITLE OF INVENTION:  4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
;  TITLE OF INVENTION:  THERMO-TOLERANCE
;  FILE REFERENCE: 4910-8
;  CURRENT APPLICATION NUMBER: US/09/654,652A
;  CURRENT FILING DATE:  2000-09-05
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 267
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme
;   OTHER INFORMATION: with enhanced activity and thermal stability
US-09-654-652A-2

  Query Match             100.0%;  Score 1333;  DB 20;  Length 267;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-130;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60

Qy     61 VLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVR 120

Qy    121 KTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 KTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS 180

Qy    181 DFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQV 240

Qy    241 PRDDEPAP 248
          ||||||||
Db    241 PRDDEPAP 248


RESULT 3
US-09-654-652A-3
; Sequence 3, Application US/09654652A
; GENERAL INFORMATION:
;  APPLICANT: SHYUR, LIE-FEN
;  APPLICANT:  CHEN, JUI-LIN
;  APPLICANT:  YANG, NING-SUN
;  TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1,
;  TITLE OF INVENTION:  4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
;  TITLE OF INVENTION:  THERMO-TOLERANCE
;  FILE REFERENCE: 4910-8
;  CURRENT APPLICATION NUMBER: US/09/654,652A
;  CURRENT FILING DATE:  2000-09-05
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Fibrobacter succinogenes
US-09-654-652A-3

  Query Match             99.6%;  Score 1328;  DB 20;  Length 349;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-129;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84

Qy     62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144

Qy    122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204

Qy    182 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 264

Qy    242 RDDEPAP 248
          |||||||
Db    265 RDDEPAP 271


RESULT 4
US-09-791-537-30441
; Sequence 30441, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30441
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Fibrobacter succinogenes
US-09-791-537-30441

  Query Match             99.6%;  Score 1328;  DB 21;  Length 349;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-129;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     25 VSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEV 84

Qy     62 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 LGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK 144
```

```
Qy    122 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 TEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSD 204

Qy    182 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 FTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVP 264

Qy    242 RDDEPAP 248
          |||||||
Db    265 RDDEPAP 271


RESULT 5
US-09-791-537-145019
; Sequence 145019, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 145019
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: pdb 1CPM
US-09-791-537-145019

  Query Match             20.4%;  Score 272.5;  DB 21;  Length 214;
  Best Local Similarity   33.8%;  Pred. No. 1.9e-19;
  Matches   77;  Conservative   33;  Mismatches   93;  Indels   25;  Gaps   12;

Qy      7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
          |  ||  :     || :|  || |  :| ||| | |      | |  | |:||| |||:
Db      1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 57

Qy     67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
          | |  |   |     ||  ::   | : |||  :| | |::| |||  :: |
Db     58 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 110

Qy    127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTL 184
           :|:  | | :  |||:      |:| :: :  ||:   :||   |||  |  ||| |
Db    111 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYTSNQ-TGGSFF-- 163

Qy    185 DWTDNFDTFDGSRWGKGDWTFDGNRVDLT--DKNI-YSRDGMLILALT 229
             : |::::   | | |   :|   : |    |: :: || | | ||
Db    164 ---EPFNSYNSGTWEKADGYSNGGVFNCTWRANNVNFTNDGKLKLGLT 208


RESULT 6
US-09-791-537-106873
; Sequence 106873, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 106873
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Brevibacillus brevis
US-09-791-537-106873

  Query Match             20.0%;  Score 266.5;  DB 21;  Length 259;
  Best Local Similarity   36.0%;  Pred. No. 1.1e-18;
  Matches   63;  Conservative   24;  Mismatches   75;  Indels   13;  Gaps    6;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          ||:::   || | :   || ||  || |   ||||| | |  |    || || |:||| |
Db     90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
          ||:    | |  |   |    :| :: :   | ::|:||  ||  : : | |:|: |
Db    149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy    123 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQ 175
             |:  | | :  |||       | | ||    |:: : :||   :||| |
Db    206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTPLQ 253


RESULT 7
US-09-791-537-32335
; Sequence 32335, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32335
;   LENGTH: 252
;   TYPE: PRT
;   ORGANISM: Brevibacillus brevis
US-09-791-537-32335

  Query Match             19.9%;  Score 265.5;  DB 21;  Length 252;
  Best Local Similarity   35.8%;  Pred. No. 1.3e-18;
  Matches   62;  Conservative   24;  Mismatches   74;  Indels   13;  Gaps    6;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          ||:::   || | :   || ||  || |   ||||| | |  |    || || |:||| |
Db     90 SARNYKAGELRTNDFYHYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFL 148

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
          ||:    | |  |   |    :| :: :   | ::|:||  ||  : : | |:|: |
Db    149 GKDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA 205

Qy    123 EGGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             |:  | | :  |||       | | ||    |:: : :||   :||| |
Db    206 ----TENIPQTPQKIMMNLWPGVGVDGWTGVFDGDNTPVYSYYDWV---RYTP 251


RESULT 8
US-09-463-862A-1
; Sequence 1, Application US/09463862A
; GENERAL INFORMATION:
;  APPLICANT: Hillen, Wolfgang
;  APPLICANT:  Maurer, Karl-Heinz
;  TITLE OF INVENTION: New Beta-Glucanase from a Bacillus
;  FILE REFERENCE: H2081PCT/US
;  CURRENT APPLICATION NUMBER: US/09/463,862A
;  CURRENT FILING DATE:  2000-05-01
;  PRIOR APPLICATION NUMBER: DE 197 32 751.6
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: PCT/EP98/04564
;  PRIOR FILING DATE: 1998-07-21
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
```

```
;   LENGTH: 308
;   TYPE: PRT
;   ORGANISM: Bacillus alkalophilus DSM 9956
US-09-463-862A-1

  Query Match             19.3%;  Score 257.5;  DB 18;  Length 308;
  Best Local Similarity   36.5%;  Pred. No. 1.2e-17;
  Matches   62;  Conservative   22;  Mismatches   73;  Indels   13;  Gaps    6;

Qy       6 DFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN 65
           ::   || | :  ||| ||  || | ::|||||:| |  |    |  || |:||| |||:
Db      96 EYKAGELRTNQFYQYGLFEVNMKPAKSTGTVSSLFTY-TGPWDWDNDPWDEIDIEFLGKD 154

Qy      66 PGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGG 125
               | |  |    |    :| :| :   | ::|:||  || |  : | |:|: |
Db     155 TTRVQFNYFTNGVG---NNEHYHELGFDASESFNTYAFEWRPESISWYVNGELVYTA--- 208

Qy     126 QVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             |:  | | :  |||      | | ||    |:    :||   :|||
Db     209 -TENIPQTPQKIMMNLWPGIGVDGWTGVFDGEDTPVVTEYDWV---RYTP 254


RESULT 9
US-09-791-537-6910
; Sequence 6910, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6910
;   LENGTH: 208
;   TYPE: PRT
;   ORGANISM: pdb 1CPN
US-09-791-537-6910

  Query Match             19.2%;  Score 256.5;  DB 21;  Length 208;
  Best Local Similarity   33.8%;  Pred. No. 8.8e-18;
  Matches   77;  Conservative   29;  Mismatches   95;  Indels   27;  Gaps   13;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |  ||  :     || :|  || |  :| ||| | |      | |  | |:||| |||:
Db       1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 57

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
               | |  |    |    ::   | : ||||  :| | |::| |||  :: |
Db      58 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 110

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTL 184
           :|:  | | :  |||:      |:| :: :  ||:   :||   |||       || |
Db     111 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT----SNGSVF-- 160

Qy     185 DWTDNFDTFDGSRWGKGDWTFDGNRVDLT--DKNI-YSRDGMLILALT 229
           |      |: | | | |   :|   : |    |: :: || | | ||
Db     161 -WEPK-SYFNPSTWEKADGYSNGGVFNCTWRANNVNFTNDGKLKLGLT 206


RESULT 10
US-09-791-537-132244
; Sequence 132244, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132244
;   LENGTH: 394
;   TYPE: PRT
;   ORGANISM: pdb 1AXKA
US-09-791-537-132244

  Query Match             18.8%;  Score 250.5;  DB 21;  Length 394;
  Best Local Similarity   29.6%;  Pred. No. 9.3e-17;
  Matches   68;  Conservative   33;  Mismatches   90;  Indels   39;  Gaps   10;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |  ||  :     || :|  || |  :| ||| | |      | |  | |:||| |||:
Db       1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 57

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
               | |  |    |    ::   | : ||||  :| | |::| |||  :: |
Db      58 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 110

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKV---------YKYTPGQ 175
           :|:  | | :  |||:      |:| :: :  ||:   :|||            :| |
Db     111 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWVKYTSNASTDYWQNWTDGG 169

Qy     176 G------EGGSDFTLDWTDNFDTFDGSRWGKGD----------WTFDGN 208
           |      | :::::|::  :   |  |  |             |  :||
Db     170 GIVNAVNGSGGNYSVNWSNTGNFVVGKGWTTGSPFRTINYNAGVWAPNGN 219


RESULT 11
US-09-791-537-117361
; Sequence 117361, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 117361
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: pdb 1AJKA
US-09-791-537-117361

  Query Match             18.6%;  Score 247.5;  DB 21;  Length 214;
  Best Local Similarity   32.9%;  Pred. No. 8e-17;
  Matches   69;  Conservative   32;  Mismatches   84;  Indels   25;  Gaps   12;

Qy      33 SGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSP 92
           :| ||| | |      | |  | |:||| |||:    | |  |    |      ||  ::
Db       2 TGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDTTKVQFNYYTNGVGGH---EKVISLGF 55

Qy      93 AADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVG 150
           | : ||||  :| | |::| |||  :: |     :|:  | | :  |||:      |:|
Db      56 DASKGFHTYAFDWQPGYIKWYVDG-VLKHT---ATANIPSTPGKIMMNLWNGTGVDDWLG 111

Qy     151 QFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRV 210
           :: :  ||:   :||   |||  |  ||| |     : |:::` | | |   :|
Db     112 SYNGAN-PLYAEYDWV---KYTSNQ-TGGSFF-----EPFNSYNSGTWEKADGYSNGGVF 161

Qy     211 DLT--DKNI-YSRDGMLILALTRKGQESFN 237
           : |    |: :: || | | ||      |:
Db     162 NCTWRANNVNFTNDGKLKLGLTSSAYNKFD 191


RESULT 12
```

```
;  FILE REFERENCE: 5914-090-999
;  CURRENT APPLICATION NUMBER: US/10/042,417
;  CURRENT FILING DATE:  2002-01-07
;  PRIOR APPLICATION NUMBER: 60/260,179
;  PRIOR FILING DATE: 2001-01-5
;  NUMBER OF SEQ ID NOS: 89
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 422
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-042-417-4

  Query Match             5.8%;  Score 83.5;  DB 12;  Length 422;
  Best Local Similarity   22.1%;  Pred. No. 3.9;
  Matches   30;  Conservative   21;  Mismatches   48;  Indels   37;  Gaps    4;

Qy    103 LEWTPNYVRWTVDGQEVRKTEGGQVSNL------------------TGTQGLRFNLWSSE 144
          | |   |::  :  :::   |  : |:|                  ||:  |   ||
Db    117 LHWKKVYLKAILRMKQLEDHEAFETSSLIGHSARVYALYYKDGLLCTGSDDLSAKLWDVS 176

Qy    145 SAAWV----------GQFDESKLPLFQFINWVKVYKYTPG------QGEGGSDFTLDWTD 188
          :   |           :||| ||    | | |  :::: |      :|  |: |:::|: |
Db    177 TGQCVYGIQTHTCAAVKFDEQKLVTGSFDNTVACWEWSSGARTQHFRGHTGAVFSVDYND 236

Qy    189 NFDTFDGSRWGKGDWT 204
            |       |  |:|
Db    237 ELDILVS---GSADFT 249


RESULT 13
US-09-841-132-192
; Sequence 192, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 192
;   LENGTH: 848
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-841-132-192

  Query Match             5.8%;  Score 83.5;  DB 10;  Length 848;
  Best Local Similarity   20.4%;  Pred. No. 9.5;
  Matches   50;  Conservative   34;  Mismatches   82;  Indels   79;  Gaps   12;

Qy      8 SGAELYTLEEVQYG-------------KFEARMKMAAASGTVSSMFL-----YQNGSEIA 49
          :| ::  :|:   |              ||:| :      |:| : ||     ::    :
Db    599 AGVDIQLMEDFVLGVSGAAFLGKMDSQKFDAEVSRKGVVGSVYTGFLAGSWFFKGQYSLG 658

Qy     50 DGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNY 109
          : :  ::    |||::  |: |  :     |    |    | |        | | : | |
Db    659 ETQNDMKTRYGVLGESSASWTSRGVL----ADALVEYRSLVGPVRPT---FYALHFNP-Y 710

Qy    110 VR----------WTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPL 159
          |           :|  |:| |  |   ::|:|   |::|  |      |   ||| |
Db    711 VEVSYASMKFPGFTEQGREARSFEDASLTNITIPLGMKFEL-----AFIKGQFSE----- 760

Qy    160 FQFIN-------WVKVYKYTPGQG----EGGSDF-----------------TLDWTDNF 190
            :|       | : |:   |      | | |:                    :|:  |
Db    761 ---VNSLGISYAW-EAYRKVEGGAVQLLEAGFDWEGAPMDLPRQELRVALENNTEWSSYF 816

Qy    191 DTFDG 195
            |  |
Db    817 STVLG 821


RESULT 14
US-09-841-132-178
; Sequence 178, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 178
;   LENGTH: 1530
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-841-132-178

  Query Match             5.8%;  Score 83.5;  DB 10;  Length 1530;
  Best Local Similarity   20.4%;  Pred. No. 20;
  Matches   50;  Conservative   34;  Mismatches   82;  Indels   79;  Gaps   12;

Qy      8 SGAELYTLEEVQYG-------------KFEARMKMAAASGTVSSMFL-----YQNGSEIA 49
          :| ::  :|:   |              ||:| :      |:| : ||     ::    :
Db   1281 AGVDIQLMEDFVLGVSGAAFLGKMDSQKFDAEVSRKGVVGSVYTGFLAGSWFFKGQYSLG 1340

Qy     50 DGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNY 109
          : :  ::    |||::  |: |  :     |    |    | |        | | : | |
Db   1341 ETQNDMKTRYGVLGESSASWTSRGVL----ADALVEYRSLVGPVRPT---FYALHFNP-Y 1392

Qy    110 VR----------WTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPL 159
          |           :|  |:| |  |   ::|:|   |::| |      |   ||| |
Db   1393 VEVSYASMKFPGFTEQGREARSFEDASLTNITIPLGMKFEL-----AFIKGQFSE----- 1442

Qy    160 FQFIN-------WVKVYKYTPGQG----EGGSDF-----------------TLDWTDNF 190
            :|       | : |:   |      | | |:                    :|:  |
Db   1443 ---VNSLGISYAW-EAYRKVEGGAVQLLEAGFDWEGAPMDLPRQELRVALENNTEWSSYF 1498

Qy    191 DTFDG 195
            |  |
Db   1499 STVLG 1503


RESULT 15
US-10-007-693-98
; Sequence 98, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
;  TITLE OF INVENTION:  AND DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.515C2
;  CURRENT APPLICATION NUMBER: US/10/007,693
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 98
;   LENGTH: 1531
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-98

  Query Match             5.8%;  Score 83.5;  DB 12;  Length 1531;
  Best Local Similarity   20.4%;  Pred. No. 20;
  Matches   50;  Conservative   34;  Mismatches   82;  Indels   79;  Gaps   12;

Qy      8 SGAELYTLEEVQYG-------------KFEARMKMAAASGTVSSMFL-----YQNGSEIA 49
```

```
          |    :|          ::||::||  ||
Db    292 DTGASYISG----------STSSIEKLMEAL 312


RESULT 10
US-09-738-363-4
; Sequence 4, Application US/09738363
; Patent No. US20010010932A1
;   GENERAL INFORMATION:
;        APPLICANT: Schnepf, Harry E.
;                   Schwab, George E.
;                   Payne, Jewel M.
;                   Narva, Kenneth E.
;                   Foncerrada, Luis
;        TITLE OF INVENTION: Nematicidal Proteins
;        NUMBER OF SEQUENCES: 40
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Jay M. Sanders
;             STREET: 2421 N.W. 41st Street, Suite A-1
;             CITY: Gainesville
;             STATE: FL
;             COUNTRY: USA
;             ZIP: 32606
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/738,363
;             FILING DATE: 15-Dec-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/076,137
;             FILING DATE: 12-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sanders, Jay
;             REGISTRATION NUMBER: 39,355
;             REFERENCE/DOCKET NUMBER: MA-20CCCD3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 352-375-8100
;             TELEFAX: 352-372-5800
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1289 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             ORGANISM: BACILLUS THURINGIENSIS
;             INDIVIDUAL ISOLATE: PS17
;        IMMEDIATE SOURCE:
;             CLONE: E. coli NM522(pMYC 1628) NRRL B-18652
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-738-363-4

  Query Match             5.9%;  Score 85.5;  DB 10;  Length 1289;
  Best Local Similarity   23.2%;  Pred. No. 11;
  Matches   56;  Conservative   31;  Mismatches   89;  Indels   65;  Gaps   13;

Qy      33 SGTVSSMFLYQNGSEIA-----DGRPWVEVD-IEVLGKNPGSFQSNIITGKAGAQKTSEK 86
           :|:: |   |||   |         | |   |      |     ||  :    |  :: :
Db    1080 NGSIRSDISYQNIDAIVLPTLPKLRHWFMSDRFSEQGDIMAKFQGALNRAYAQLEQNTLL 1139

Qy      87 HHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSS--E 144
           |:                :| :   |||:|      :  ||    | : ||   |||
Db    1140 HNG--------------HFTKDAANWTVEG------DAHQVVLEDGKRVLRLPDWSSSVS 1179

Qy     145 SAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLD----------WTDNFDTFD 194
                :  ||  |   :| :          |||||    ||:          | :|  |

Db    1180 QTIEIENFDPDK--EYQLV--------FHGQGEG--TVTLEHGEETKYIETHTHHFANFT 1227

Qy     195 GSRWGKGDWTFDGNRVDLTDKNIYSRDGMLI---LALTRKGQESFNGQVPRDDEPAPNSS 251
           |:  :    ||: |:| :|    | | ||  :    :||          :| ||: :  ::
Db    1228 TSQ--RQGLTFESNKVTVT---ISSEDGEFLVDNIALV-------EAPLPTDDQNSEGNT 1275

Qy     252 S 252
           :
Db    1276 A 1276


RESULT 11
US-09-738-626-6377
; Sequence 6377, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6377
;   LENGTH: 714
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-6377

  Query Match             5.8%;  Score 84;  DB 9;  Length 714;
  Best Local Similarity   22.5%;  Pred. No. 6.9;
  Matches   51;  Conservative   36;  Mismatches  102;  Indels   38;  Gaps   11;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAAS-----GTVSSMFLYQNGSEIA--DGRPWVEVDI 59
           :| | :: | |:  | |   ::  | |      |   ::::  | :  |  ||||: |:||
Db     436 YSSARVH-LPEIPAGNFRLTVRARAQSEELVDGVRPAIWMQNNTNFCADNDGRPYGELDI 494

Qy      60 -EVLGKNPGSFQSNIITGKAG---AQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVD 115
            |       :  | :  |-||     |  :     |   |  :| :|:|     : :|:|
Db     495 TEFYSSRVNTQYSAVHLGCAGNRPEMKLRQMEMEESMFGD--WHDWGVEVFDGQIVFTID 552

Qy     116 GQEV----RKTEGGQVSNLTG-TQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYK 170
           |: |    :   |  |:      :   | |   |    :||      |    :| :
Db     553 GKAVTSSGKDVFGNSVTPAAAPLRPAHFKLSEEEYREVIGQ------PWHLILNTM---- 602

Qy     171 YTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 217
                 : | |  :   || : |   |     :  |   ||:   ::
Db     603 ----VEQSGKDSWITAVDNNEAFPEHR-----FQIDHVAVDIESDSV 640


RESULT 12
US-10-042-417-4
; Sequence 4, Application US/10042417
; Patent No. US20020123082A1
; GENERAL INFORMATION:
;  APPLICANT: Pagano, M.
;  TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
;  TITLE OF INVENTION:  PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
```

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial
;   OTHER INFORMATION: Sequence:Pelb/4AF/myc/6HIS
US-09-818-247-22

  Query Match             6.2%;  Score 89;  DB 10;  Length 288;
  Best Local Similarity   19.2%;  Pred. No. 0.75;
  Matches   54;  Conservative   42;  Mismatches  113;  Indels   72;  Gaps   11;

Qy      26 RMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSE 85
           |:  ||:  | ||  :          |  ||     : |    :: :: : |:    : :
Db      41 RLSCAASGFTFSSYAMSWVRQAPGKGLEWVSA---ISGSGGSTYYADSVKGRFTISRDNS 97

Qy      86 KHHAVSP-----AADQAFHTYGLEWTPN--YVRWTVDGQEVRKTEGGQVSNLTGTQGLRF 138
           |:          | | | :     :| |  | :    |  |  : ||  |   |: |
Db      98 KNTLYLQMNSLRAEDTAVYYCARSFTVNSGYFQHWGQGTLVTVSSGGGGSGGGGSGGGGS 157

Qy     139 NLWSSESAAWVGQFDESKLPLF-----QFINWVKVYKYTPGQ------------------ 175
             :  ::| : :       :: :          :|:  |:  ||:
Db     158 EIVLTQSPSTLSASIGDRVTITCRASEGIYHWLAWYQQKPGKAPKLLIYKASSLASGVPS 217

Qy     176 ---GEG-GSDFTLDWT----DNFDTFDGSRWGKGDWTF-DGNRVDLTDKNIYSRDGMLIL 226
             | | |:|||  :    ::| |:     :     ||  | :||:  |   :   ::
Db     218 RFSGSGSGTDFTLTISSLQPEDFATYYCQHYDSTPPTFGQGTKVDI--KRAAAEQKLI-- 273

Qy     227 ALTRKGQESFNGQVPRDDEPAPNSSSVDKLAAALEHHHHHH 267
                :|  ||                          ||||||
Db     274 -----SEEDLNGAA--------------------HHHHHH 288


RESULT 8
US-09-738-363-2
; Sequence 2, Application US/09738363
; Patent No. US20010010932A1
;   GENERAL INFORMATION:
;        APPLICANT: Schnepf, Harry E.
;                   Schwab, George E.
;                   Payne, Jewel M.
;                   Narva, Kenneth E.
;                   Foncerrada, Luis
;        TITLE OF INVENTION: Nematicidal Proteins
;        NUMBER OF SEQUENCES: 40
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Jay M. Sanders
;             STREET: 2421 N.W. 41st Street, Suite A-1
;             CITY: Gainesville
;             STATE: FL
;             COUNTRY: USA
;             ZIP: 32606
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/738,363
;             FILING DATE: 15-Dec-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/076,137
;             FILING DATE: 12-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sanders, Jay
;             REGISTRATION NUMBER: 39,355
;             REFERENCE/DOCKET NUMBER: MA-20CCCD3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 352-375-8100
;             TELEFAX: 352-372-5800
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1385 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             ORGANISM: BACILLUS THURINGIENSIS
;             INDIVIDUAL ISOLATE: PS17
;        IMMEDIATE SOURCE:
;             CLONE: E. coli NM522(pMYC 1627) NRRL B-18651
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-738-363-2

  Query Match             6.1%;  Score 87.5;  DB 10;  Length 1385;
  Best Local Similarity   23.3%;  Pred. No. 7.7;
  Matches   56;  Conservative   33;  Mismatches   88;  Indels   63;  Gaps   13;

Qy      33 SGTVSSMFLYQNGSEIA-----DGRPWVEVD-IEVLGKNPGSFQSNIITGKAGAQKTSEK 86
           :|:: |   |||   |        | |   |     |     ||  :    |  ::::
Db    1176 NGSIRSDISYQNIDAIVLPTLPKLRHWFMSDRFSEQGDIMAKFQGALNRAYAQLEQSTLL 1235

Qy      87 HHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESA 146
           |:              :| :   ||::|      :  |::   | : ||   |||
Db    1236 HNG-------------HFTKDAANWTIEG-----DAHQITLEDGRRVLRLPDWSSS-- 1273

Qy     147 AWVGQFDESKLPLFQFINWVKVYKYT-PGQGEGGSDFTLD----------WTDNFDTFDG 195
             | |  |       :  |  | |     |||||    ||:          | :|  |
Db    1274 --VSQMIE-----IENFNPDKEYNLVFHGQGEG--TVTLEHGEETKYIETHTHHFANFTT 1324

Qy     196 SRWGKGDWTFDGNRVDLTDKNIYSRDGMLI---LALTRKGQESFNGQVPRDDEPAPNSSS 252
           |:  :   ||: |:| :|   | | ||  :   :||           :| ||: :  :::
Db    1325 SQ--RQGLTFESNKVTVT---ISSEDGEFLVDNIALV-------EAPLPTDDQNSEGNTA 1372


RESULT 9
US-09-215-450-26
; Sequence 26, Application US/09215450
; Patent No. US20020068278A1
; GENERAL INFORMATION:
;  APPLICANT: Giese, Klaus
;  APPLICANT: Xin, Hong
;  TITLE OF INVENTION: METASTATIC BREAST AND COLON CANCER REGULATED GENES
;  FILE REFERENCE: 1451.100 / 210030.447
;  CURRENT APPLICATION NUMBER: US/09/215,450
;  CURRENT FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 26
;   LENGTH: 406
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-215-450-26

  Query Match             6.0%;  Score 86.5;  DB 10;  Length 406;
  Best Local Similarity   23.2%;  Pred. No. 2;
  Matches   63;  Conservative   34;  Mismatches  115;  Indels   59;  Gaps   13;

Qy      27 MKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEK 86
           ||       | ||: |    :   | : : |: |   |  | :|:    || :     | |
Db      64 MKRLTLGNTTSSVIL----TNYMDTQYYGEIGI---GTPPQTFKVVFDTGSSNVWVPSSK 116

Qy      87 ---------HHAVSPAADQAFHTY-GLEWTPNYVRWTVDG---QEVRKTEGGQVSNLTG- 132
                   :| :  |:| : : : | | |   |   || |   |::    |  |: : |
Db     117 CSRLYTACVYHKLFDASDSSSYKHNGTELTLRYSTGTVSGFLSQDIITVGGITVTQMFGE 176

Qy     133 ---TQGLRFNLWSSESAAWVGQFDES---KLPLFQFINWVKVYK------YTPGQGE--- 177
                | | |   :    :|  :::     |:|  |    | |      |     |
Db     177 VTEMPALPFMLAEFDGVVGMGFIEQAIGRVTPIFDNIISQGVLKEDVFSFYYNRDSENSQ 236

Qy     178 --------GGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALT 229
                   ||||    :  ||   :  :  | |      | :    :   ||   |||
Db     237 SLGGQIVLGGSD-PQHYEGNFHYINLIK--TGVWQIQMKGVSVGSSTLLCEDG--CLALV 291

Qy     230 RKGQESFNGQVPRDDEPAPNSSSVDKLAAAL 260
```

```
Db    168 QFDWYEGHQDDIYDMDLNLHAVVKENGQGVWKRPKMYPQEQLNKWRAMDPSKD--FHIYG 225

Qy    103 LEWTPNYVRWTVDGQEVRKTEG---GQVSNLTGTQGLR---FNLWSSESAAWVGQFDE-- 154

Db    226 CEVNQNEIIWYVDGVEVARKPNKYWHRPMNVTLSLGLRKPFVKFFDNKNNAINPETDAKA 285

Qy    155 ----SKLPLFQFINWVKVYKYTPG 174

Db    286 REKLSDIPTSMYVDYVRVWEKSAG 309


RESULT 4
US-09-731-221-79
; Sequence 79, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
;  APPLICANT: Caplan, Michael
;  TITLE OF INVENTION: Passive Desensitization
;  FILE REFERENCE: 2002834-0103
;  CURRENT APPLICATION NUMBER: US/09/731,221
;  CURRENT FILING DATE:  2001-12-06
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
;   LENGTH: 526
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Arachis
;   OTHER INFORMATION: Hypogaea
US-09-731-221-79

  Query Match             7.0%;  Score 100.5;  DB 10;  Length 526;
  Best Local Similarity   54.1%;  Pred. No. 0.14;
  Matches   20;  Conservative    4;  Mismatches   12;  Indels    1;  Gaps    1;

Qy    231 KGQESFNGQVPRDDEPAPNSSSVDKLAAALEHHHHHH 267

Db    491 KNNNPFKFFVP-PSQQSPRAVAVDKLAAALEHHHHHH 526


RESULT 5
US-09-804-626-2
; Sequence 2, Application US/09804626
; Patent No. US20020128190A1
; GENERAL INFORMATION:
;  APPLICANT: Lobel, Leslie
;  APPLICANT:  Lustbader, Joyce
;  TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR DOMAIN OF
;  TITLE OF INVENTION:  GONADOTROPIN RECEPTOR
;  FILE REFERENCE: 0575/62259/JPW/SHS
;  CURRENT APPLICATION NUMBER: US/09/804,626
;  CURRENT FILING DATE:  2001-03-09
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-804-626-2

  Query Match             6.7%;  Score 97;  DB 10;  Length 518;
  Best Local Similarity   22.3%;  Pred. No. 0.3;
  Matches   60;  Conservative   33;  Mismatches   78;  Indels   98;  Gaps   13;

Qy     20 YGK-FEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKA 78

Db    327 YGNGFEEVQSHAFNGTTLTSLELKEN-----------VHLEKM--HNGAFR-------- 364

Qy     79 GAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRF 138

Db    365 GA--TGPKTLDISSTKLQALPSYGLE--------------------SIQRLIAT----- 396

Qy    139 NLWSSESAAWVGQFDESKLPLFQ-FINWVK----------VYKYTPGQGEGGSDFTLDWT 187

Db    397 -----------SSYSLKKLPSRETFVNLLEATLTYPSHCCAFRNLPTKEQ---NFSHSIS 442

Qy    188 DNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVPRDDEPA 247

Db    443 ENFSKQCESTVRK-----------VNNKTLYS--SMLAESELSGWDYEYGFCLPKTPRCA 489

Qy    248 PNSSS---------VDKLAAALEHHHHHH 267

Db    490 PEPDAFNPCEDIMGVDKLAAALEHHHHHH 518


RESULT 6
US-09-804-626-4
; Sequence 4, Application US/09804626
; Patent No. US20020128190A1
; GENERAL INFORMATION:
;  APPLICANT: Lobel, Leslie
;  APPLICANT:  Lustbader, Joyce
;  TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR DOMAIN
;  TITLE OF INVENTION:  GONADOTROPIN RECEPTOR
;  FILE REFERENCE: 0575/62259/JPW/SHS
;  CURRENT APPLICATION NUMBER: US/09/804,626
;  CURRENT FILING DATE:  2001-03-09
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 516
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-804-626-4

  Query Match             6.7%;  Score 96.5;  DB 10;  Length 516;
  Best Local Similarity   38.2%;  Pred. No. 0.33;
  Matches   26;  Conservative    5;  Mismatches   30;  Indels    7;  Gaps    2;

Qy    207 GNRVDLTDKNIYSRDGMLILALTRKGQESFNGQV-----PRDD--EPAPNSSSVDKLAAA 259

Db    449 GQRSSLAEDNESSYSRGFDMTYTEFDYDLCNEVVDVTCSPKPDAFNPCEDIMGVDKLAAA 508

Qy    260 LEHHHHHH 267

Db    509 LEHHHHHH 516


RESULT 7
US-09-818-247-22
; Sequence 22, Application US/09818247
; Patent No. US20020102657A1
; GENERAL INFORMATION:
;  APPLICANT: Mostov, Keith E.
;  APPLICANT:  Chapin, Steven J.
;  APPLICANT:  Richman-Eisenstat, Janice
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Ligands Directed to the No. US20020102657A1-Secretory Componen
;  TITLE OF INVENTION:  No. US20020102657A1-Stalk Region of pIgR and Methods of Use T
;  FILE REFERENCE: 18062E-000910US
;  CURRENT APPLICATION NUMBER: US/09/818,247
;  CURRENT FILING DATE:  2001-03-26
;  PRIOR APPLICATION NUMBER: WO PCT/US01/09699
;  PRIOR FILING DATE: 2001-03-26
;  PRIOR APPLICATION NUMBER: US 60/192,197
;  PRIOR FILING DATE: 2000-03-27
;  PRIOR APPLICATION NUMBER: US 60/192,198
;  PRIOR FILING DATE: 2000-03-27
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
;   LENGTH: 288
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
```

```
Qy    134 ----------QGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFT 183
                  |:  ::|   | |  |     ::|    |||              : ||
Db    175 PGTKYLNQQAMGVYISIWDGSSWATQG----GRVP----INW-----------ASAPFT 214

Qy    184 LDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKN 216
            :              |:  :|  ||  | |
Db    215 ATYQ--------------DFALNGCVVDPNDPN 233


RESULT 2
US-09-988-200-6
; Sequence 6, Application US/09988200
; Patent No. US20020094553A1
;   GENERAL INFORMATION:
;        APPLICANT: BARBEYRON, Tristan
;                   POTIN, Philippe
;                   RICHARD, Christophe
;                   HENRISSAT, Bernard
;                   YVIN, Jean-Claude
;                   KLOAREG, Bernard
;        TITLE OF INVENTION: Glycolyse hydrolase genes and their
;                            use for producing enzymes for the biodegradation of
;                            carrageenans
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DENNISON, MESEROLE, SCHEINER & SCHULTZ
;             STREET: 612 Crystal Square 4, 1745 Jefferson Davis
;                     Highway
;             CITY: ARLINGTON
;             STATE: VIRGINIA
;             COUNTRY: U.S.A.
;             ZIP: 22202
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/988,200
;             FILING DATE: 19-No. US20020094553A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/269,731
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: FR 96 12204
;             FILING DATE: 07-OCT-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: IRA SCHULTZ
;             REGISTRATION NUMBER: <Unknown>
;             REFERENCE/DOCKET NUMBER: <Unknown>
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 412-1155
;             TELEFAX: (703) 412-1161
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 620 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-988-200-6

  Query Match             8.1%;  Score 117;  DB 10;  Length 620;
  Best Local Similarity   27.3%;  Pred. No. 0.0055;
  Matches   54;  Conservative   23;  Mismatches   73;  Indels   48;  Gaps   10;

Qy     20 YGKFEARMKMAAASGTVSSMF-LYQ--NGSEIADGR-PWVEVDI---------------- 59
          || :|||:| |:    ||  | :|   : |   :|   : |:|:
Db    289 YGYYEARIKGASTFPGVSPAFWMYSTIDRSLTKEGDVQYSEIDVVELTQKSAVRESDHDL 348

Qy     60 -EVLGKN-------PGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVR 111
           :: ||       ||||      |          |    |  | |||||:  | : :
```

```
Db    349 HNIVVKNGKPTWMRPGSFPQTNHNG---------YHLPFDPRND--FHTYGVNVTKDKIT 397

Qy    112 WTVDGQEVRKTEG---GQVSNLTGTQGLR--FNLWSSE----SAAWVGQFDESKLPLFQF 162
          | |||: | : :     :  ||| :|||     |       ||    :   : : :
Db    398 WYVDGEIVGEKDNLYWHRQMNLTLSQGLRAPHTQWKCNQFYPSANKSAEGFPTSMEVDYV 457

Qy    163 INWVKVYKYTPGQGEGGS 180
            ||||       |.| |
Db    458 RTWVKVGNNNSAPGEGQS 475


RESULT 3
US-09-988-200-8
; Sequence 8, Application US/09988200
; Patent No. US20020094553A1
;   GENERAL INFORMATION:
;        APPLICANT: BARBEYRON, Tristan
;                   POTIN, Philippe
;                   RICHARD, Christophe
;                   HENRISSAT, Bernard
;                   YVIN, Jean-Claude
;                   KLOAREG, Bernard
;        TITLE OF INVENTION: Glycolyse hydrolase genes and their
;                            use for producing enzymes for the biodegradation of
;                            carrageenans
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DENNISON, MESEROLE, SCHEINER & SCHULTZ
;             STREET: 612 Crystal Square 4, 1745 Jefferson Davis
;                     Highway
;             CITY: ARLINGTON
;             STATE: VIRGINIA
;             COUNTRY: U.S.A.
;             ZIP: 22202
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/988,200
;             FILING DATE: 19-No. US20020094553A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/269,731
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: FR 96 12204
;             FILING DATE: 07-OCT-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: IRA SCHULTZ
;             REGISTRATION NUMBER: <Unknown>
;             REFERENCE/DOCKET NUMBER: <Unknown>
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 412-1155
;             TELEFAX: (703) 412-1161
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 545 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-988-200-8

  Query Match             7.9%;  Score 113;  DB 10;  Length 545;
  Best Local Similarity   23.0%;  Pred. No. 0.011;
  Matches   47;  Conservative   41;  Mismatches   78;  Indels   38;  Gaps   10;

Qy      7 FSGAELYTLEEVQYGKFEARMKMA-AASGTVSSMFLYQN-GSEIADGRP-WVEVDIEVL- 62
          |:     : ::  || |||::: |    |   | :|| :    :|:|   : |:|:  |
Db    108 FTSGIFKSYQKFTYGYFEAKIQGADIGEGVCPSFWLYSDFDYSVANGETVYSEIDVVELQ 167

Qy     63 ------GKNPGSFQSNI----ITGKAG----------AQKTSEKHHAVSPAADQAFHTYG 102
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:16 ; Search time 6.79861 Seconds
(without alignments)
761.932 Million cell updates/sec

Title: US-09-654-652A-2
Perfect score: 1439
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....PNSSSVDKLAAALEHHHHHH 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	120	8.3	287	10	US-09-734-569-152	Sequence 152, App
2	117	8.1	620	10	US-09-988-200-6	Sequence 6, Appli
3	113	7.9	545	10	US-09-988-200-8	Sequence 8, Appli
4	100.5	7.0	526	10	US-09-731-221-79	Sequence 79, Appl
5	97	6.7	518	10	US-09-804-626-2	Sequence 2, Appli
6	96.5	6.7	516	10	US-09-804-626-4	Sequence 4, Appli
7	89	6.2	288	10	US-09-818-247-22	Sequence 22, Appl
8	87.5	6.1	1385	10	US-09-738-363-2	Sequence 2, Appli
9	86.5	6.0	406	10	US-09-215-450-26	Sequence 26, Appl
10	85.5	5.9	1289	10	US-09-738-363-4	Sequence 4, Appli
11	84	5.8	714	9	US-09-738-626-6377	Sequence 6377, Ap
12	83.5	5.8	422	12	US-10-042-417-4	Sequence 4, Appli
13	83.5	5.8	848	10	US-09-841-132-192	Sequence 192, App
14	83.5	5.8	1530	10	US-09-841-132-178	Sequence 178, App
15	83.5	5.8	1531	12	US-10-007-693-98	Sequence 98, Appl
16	82.5	5.7	381	10	US-09-822-698A-5	Sequence 5, Appli
17	81.5	5.7	716	9	US-09-828-523A-82	Sequence 82, Appl
18	81	5.6	392	9	US-09-813-718-16	Sequence 16, Appl
19	81	5.6	415	9	US-09-813-718-14	Sequence 14, Appl
20	81	5.6	437	9	US-09-813-718-12	Sequence 12, Appl
21	81	5.6	484	9	US-09-813-718-10	Sequence 10, Appl
22	81	5.6	518	9	US-10-098-174-1	Sequence 1, Appli
23	80.5	5.6	467	10	US-09-863-547B-1	Sequence 1, Appli
24	80	5.6	156	10	US-09-734-569-28	Sequence 28, Appl
25	80	5.6	634	10	US-09-731-221-78	Sequence 78, Appl
26	78.5	5.5	201	10	US-09-988-974-5	Sequence 5, Appli
27	78.5	5.5	201	10	US-09-988-974-9	Sequence 9, Appli
28	78.5	5.5	209	10	US-09-864-761-42996	Sequence 42996, A
29	78.5	5.5	475	9	US-09-738-626-6052	Sequence 6052, Ap
30	78.5	5.5	663	10	US-09-815-242-5350	Sequence 5350, Ap
31	78.5	5.5	663	10	US-09-815-242-12138	Sequence 12138, A
32	78.5	5.5	764	9	US-10-105-695-4	Sequence 4, Appli
33	78.5	5.5	764	9	US-10-105-694-4	Sequence 4, Appli
34	78.5	5.5	764	10	US-09-747-521-4	Sequence 4, Appli
35	78.5	5.5	764	12	US-10-106-014-4	Sequence 4, Appli
36	78	5.4	166	10	US-09-731-221-77	Sequence 77, Appl
37	78	5.4	862	10	US-09-833-435A-1	Sequence 1, Appli
38	77.5	5.4	231	10	US-09-728-911-2	Sequence 2, Appli
39	77.5	5.4	231	10	US-09-949-192-6	Sequence 6, Appli
40	77	5.4	200	10	US-09-734-569-150	Sequence 150, App
41	77	5.4	1233	9	US-09-738-626-4312	Sequence 4312, Ap
42	76.5	5.3	252	10	US-09-750-424-1	Sequence 1, Appli
43	76.5	5.3	1050	10	US-09-866-562-38	Sequence 38, Appl
44	76	5.3	1016	12	US-10-044-303-2	Sequence 2, Appli
45	75.5	5.2	165	10	US-09-925-299-865	Sequence 865, App

ALIGNMENTS

RESULT 1
US-09-734-569-152
; Sequence 152, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv
; TITLE OF INVENTION: in the synthesis of carbohydrates
; FILE REFERENCE: BASF-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734,569
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 152
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-152

Query Match 8.3%; Score 120; DB 10; Length 287;
Best Local Similarity 23.0%; Pred. No. 0.0011;
Matches 49; Conservative 30; Mismatches 70; Indels 64; Gaps 11;

```
Qy      20 YGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGS---FQSNII 74
           |    | :||    ::|||:: ::   | :        |:|:| ||   |      :|:
Db      69 YVDISAYIKMPPFDSAGTVTTFYMSSQGDQ------HYELDMEFLGNTSGQPFLLHTNVF 122

Qy      75 TGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGT- 133
           | ::  : :    |:|| || |   |: : | : || : ||       || ||
Db     123 VDGVGGRE-QQMYLGFDPSAD--FHYYRFRWSKDMVVFYVDNKPVR-----VFKNLEGTV 174
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608551; PubMed=11743194;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
DR   EMBL; AE009336; AAL44856.1; ALT_INIT.
DR   EMBL; AE008279; AAK89373.1; -.
KW   Complete proteome.
SQ   SEQUENCE   289 AA;  32368 MW;  0384C6F83320EAC9 CRC64;

  Query Match            14.3%;  Score 206;  DB 16;  Length 289;
  Best Local Similarity   35.1%;  Pred. No. 8.9e-09;
  Matches   54;  Conservative   26;  Mismatches   56;  Indels   18;  Gaps    6;

Qy       5 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
           ::|:  |: |    :|| :||||| |  ||  |: | |   :   | :|  |:| |||||
Db     115 RNFACGEIQTKGRYRYGTYEARMKAATGSGLNSAFFTYIGPT---DKKPHDEIDFEVLGK 171

Qy      65 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
           | |  | |  | | |    :||   |   ||  |: |   | |  :|: |:|: |
Db     172 NTGKVQLNQYIAAKGG----NEKLVPVEGGADAGFNDYAFVWEPQRLRYYVNGKLVH--- 224

Qy     124 GGQVSNLT----GTQGLRFNLWSSESAA-WVGQF 152
             :|:: |      | : |:|| :::   |:|:|
Db     225 --EVTDETKIPQNAQKIFFSLWGTDTLKDWMGKF 256


RESULT 15
Q98C78
ID   Q98C78      PRELIMINARY;      PRT;   293 AA.
AC   Q98C78;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Endo-1,3-1,4-beta-glycanase, ExoK.
GN   MLR5264.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003006; BAB51743.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
KW   Complete proteome.
SQ   SEQUENCE   293 AA;  32347 MW;  93BBA1672176ED3C CRC64;

  Query Match            13.9%;  Score 200;  DB 16;  Length 293;
  Best Local Similarity   33.8%;  Pred. No. 2.8e-08;
  Matches   51;  Conservative   27;  Mismatches   61;  Indels   12;  Gaps    6;

Qy       5 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
           ::|:  |: | :   || :|||:|    ||  :: | |   |   | :|| |:| |:| |
Db     112 REFACGEIQTKQRFGYGTYEARLKTDTGSGLNAAFFTYIGPS---DKQPWDEIDFEILTK 168

Qy      65 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK-T 122
           :    | |  | ||   :|  |    |    |:||: |   |  : :|| |:|| |   |
Db     169 DTSKVQVNAYIDGKGKNEKLVE----VPGGTDKAFNDYAFVWEKDSLRWYVNGQLVNTIT 224

Qy     123 EGGQVSNLTGTQGLRFNLWSSES-AAWVGQF 152
           :  ::   :  | : |:|| ||:   |:| |
Db     225 DPAKLP--SHAQKIFFSLWGSETMKGWMGAF 253


Search completed: January  9, 2003, 12:18:12
Job time : 25.7951 secs
```

```
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    237       BETA-(1,3-1,4)-GLUCANASE.
SQ   SEQUENCE   237 AA;  26989 MW;  7DEF5BCE53790470 CRC64;

  Query Match             16.0%;  Score 230;  DB 2;  Length 237;
  Best Local Similarity   33.1%;  Pred. No. 8.1e-11;
  Matches   54;  Conservative   26;  Mismatches   69;  Indels   14;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db      81 YTGGEWRSKERFGYGLFQVNMKPIKNPGVVSSFFTYTGPS---DGTKWDEIDIEFLGKDT 137

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
Db     138 TKVQFNYYT---SGQGNHEYLYNLGFDASQGFHTYGFDWQADHITWYVDGRAVYTA---- 190

Qy     127 VSNLTGTQG-LRFNLW--SSESAAWVGQFDESKLPLFQFINWV 166
Db     191 YNNIPSTPGKIMMNAWPGTHEVDSWLGAYN-GRTPLYAYYDWI 232


RESULT 12
Q97FD3
ID   Q97FD3      PRELIMINARY;      PRT;   246 AA.
AC   Q97FD3;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-1,3(4)-beta-glucanase family 16.
GN   CAC2807.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
DR   EMBL; AE007778; AAK80751.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Complete proteome.
SQ   SEQUENCE   246 AA;  27717 MW;  C0EBC302678D2FE1 CRC64;

  Query Match             15.4%;  Score 221;  DB 16;  Length 246;
  Best Local Similarity   33.7%;  Pred. No. 4.5e-10;
  Matches   58;  Conservative   19;  Mismatches   73;  Indels   22;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db      91 YAGGEYRSNNRYGYGLYRVSMKPAKHIGVDSSFFSYTGPS---DNNPWDEIDIEFLGKDT 147

Qy      67 GSFQSNIITGKAGAQKTSEKH---HAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
Db     148 TEVQFNYYTNGVG------KHEYLYKLGFDASKGFHTYGYIWEQNYIAWLVDGKEVYRA- 200

Qy     124 GGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
Db     201 ---TSNIPTHPGKVMMNLWPGIGVDSWLGAYD-GVTPVKAYYNWA---MYNP 245


RESULT 13
Q93UI1
ID   Q93UI1      PRELIMINARY;      PRT;   205 AA.
AC   Q93UI1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254960; AAK50613.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER     205    205
SQ   SEQUENCE   205 AA;  22890 MW;  E7CADAAFE02A2F51 CRC64;

  Query Match             15.2%;  Score 219;  DB 2;  Length 205;
  Best Local Similarity   36.2%;  Pred. No. 5.1e-10;
  Matches   55;  Conservative   20;  Mismatches   67;  Indels   10;  Gaps    5;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db      55 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIESL 111

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
Db     112 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHMYAFDWQPNSIKWYVDGQ-LKHT 167

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 153
Db     168 ATTQVPAAPGK--IMMNLWNGTGVDDWLGSYN 197


RESULT 14
Q8U8N5
ID   Q8U8N5      PRELIMINARY;      PRT;   289 AA.
AC   Q8U8N5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Endo-1,3-1,4-beta-glycanase.
GN   EXOK OR ATU4055 OR AGR_L_1600.
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193;
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.;
RT   "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT   C58.";
RL   Science 294:2317-2323(2001).
RN   [2]
```

```
AC   Q45691;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-beta-1,3-1,4-glucanase.
GN   BEG1.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCIB 8565;
RA   van Rensburg P., van Zyl W.H., Pretorius I.S.;
RT   "Over-expression of the Saccharomyces cerevisiae exo-beta-1,3-
RT   glucanase gene together with the Butyrivibrio fibrisolvens endo-beta-
RT   1,4-glucanase gene and the Bacillus subtilis endo-beta-1,3-1,4-
RT   glucanase gene in Saccharomyces cerevisiae.";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U60830; AAB05759.1; -.
DR   HSSP; P27051; 1GBG.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
SQ   SEQUENCE   239 AA;  26969 MW;  C4219760D13F878A CRC64;

  Query Match             16.3%;  Score 234.5;  DB 2;  Length 239;
  Best Local Similarity   33.7%;  Pred. No. 3.5e-11;
  Matches   58;  Conservative   26;  Mismatches   71;  Indels   17;  Gaps    7;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      80 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 136

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPA---ADQAFHTYGLEWTPNYVRWTVDGQEV 119
           ||:    | |  |  |      | |  ::     |  |:|||  :| || ::| || | :
Db     137 GKDTTKVQFNYYTNGA------ENHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDWQ-L 189

Qy     120 RKTEGGQVSNLTGTQGLRFNLWSSESAA-WVGQFDESKLPLFQFINWVKVYK 170
           : |   |:   |   :  |||:   : |:| ::   ||:   :||:  |
Db     190 KHTATTQIPTAPGK--IMMNLWNGTGVSDWLGSYNGVN-PLYAHYDWVRYTK 238


RESULT 9
Q8RMP0
ID   Q8RMP0      PRELIMINARY;      PRT;   242 AA.
AC   Q8RMP0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Beta-1,3-1,4-glucanase.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sun J., Li W., Gu S., Xu Z., Zhao H., Xiao J.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF490978; AAM08358.1; -.
SQ   SEQUENCE   242 AA;  27238 MW;  285FED3FE76AE69A CRC64;

  Query Match             16.2%;  Score 232.5;  DB 2;  Length 242;
  Best Local Similarity   34.9%;  Pred. No. 5.2e-11;
  Matches   59;  Conservative   24;  Mismatches   75;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |
Db      83 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFL 139

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
Db     140 GKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 195

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           |:    |   :  |||:      :| ::    ||:   :||:  |
Db     196 ATNQIPTTPGK--IMMNLWNGTGVDERLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 10
Q9FDC9
ID   Q9FDC9      PRELIMINARY;      PRT;   212 AA.
AC   Q9FDC9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-1,3-1,4-beta-glucanase (Fragment).
OS   Paenibacillus polymyxa (Bacillus polymyxa).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Paenibacillaceae; Paenibacillus.
OX   NCBI_TaxID=1406;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yao W., Wang Y., Song W., Yang K., Su Z.;
RT   "Gene cloning of an antifugal protein.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF284449; AAG02415.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
FT   NON_TER       1      1
FT   NON_TER     212    212
SQ   SEQUENCE   212 AA;  24127 MW;  7E707E7FBEF7A440 CRC64;

  Query Match             16.0%;  Score 230;  DB 2;  Length 212;
  Best Local Similarity   35.1%;  Pred. No. 7e-11;
  Matches   59;  Conservative   23;  Mismatches   70;  Indels   16;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           | | |  :    :|| :|  || |  :| ||| | |      |:|  | |:||| |||:
Db      57 FDGGEYRSKNTYRYGLYEVNMKPAKNTGIVSSFFTYTGP---ANGTQWDEIDIEFLGKDT 113

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
           | |  |   |     ||   :   |  ||||  :| | |::| |||  :: |
Db     114 TKVQFNYYTNGIGGH---EKVVDLGFDASSGFHTYAFDWQPGYIKWYVDG-VLKHT---A 166

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
           :|:  | | :  |||:     :|:| ::    ||:   :||   |||
Db     167 TTNIPKTPGQIMMNLWNGTGVDSWLGPYNGVN-PLYAEYDWV---KYT 210


RESULT 11
O07856
ID   O07856      PRELIMINARY;      PRT;   237 AA.
AC   O07856;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Beta-(1,3-1,4)-glucanase precursor.
GN   LICHENASE.
OS   Streptococcus bovis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Streptococcus.
OX   NCBI_TaxID=1315;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JB1;
RA   Ekinci M., Flint H.J.;
RT   "Isolation and overexpression of a gene encoding an extracellular
RT   beta-(1,3-1,4)-glucanase from Streptococcus bovis JB1.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z92911; CAB07443.1; -.
DR   HSSP; P23904; 1AJK.
```

```
Db    112 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 167

Qy    123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
          |:    |   :  |||:      |:| ::    ||:   :||:  |
Db    168 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 213


RESULT 5
Q9S310
ID   Q9S310      PRELIMINARY;      PRT;   802 AA.
AC   Q9S310;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Family 11 xylanase /family 16 beta (1,3-1,4) glucanase.
GN   XYND.
OS   Ruminococcus flavefaciens.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Lachnospiraceae; Ruminococcus.
OX   NCBI_TaxID=1265;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=17;
RA   Aurilia V., Martin J.C., Munro C.A., Mercer D.K., Flint H.J.;
RT   "Organisation and strain distribution of genes responsible for the
RT   utilization of xylans by the rumen cellulolytic bacterium Ruminococcus
RT   flavefaciens 17.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ132472; CAB51934.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR003305; CBM_CenC.
DR   InterPro; IPR001137; GH_11.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF02018; CBM_4_9; 2.
DR   Pfam; PF00457; Glyco_hydro_11; 2.
DR   Pfam; PF00722; Glyco_hydro_16; 2.
DR   PRINTS; PR00911; GLHYDRLASE11.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS00776; GLYCOSYL_HYDROL_F11_1; 2.
DR   PROSITE; PS00777; GLYCOSYL_HYDROL_F11_2; 2.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 2.
FT   CHAIN        2    802       FAMILY 11 XYLANASE /FAMILY 16 BETA
FT                               (1,3-1,4) GLUCANASE.
SQ   SEQUENCE   802 AA;  89019 MW;  F36BC68805FC5274 CRC64;

  Query Match             16.9%;  Score 242.5;  DB 2;  Length 802;
  Best Local Similarity   34.9%;  Pred. No. 4e-11;
  Matches   58;  Conservative   21;  Mismatches   74;  Indels   13;  Gaps    6;

Qy      7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
          :|| |  |     || :|  |:     | ||| | |   |   |  || |:|||:||||
Db    634 YSGGEFRTNNFYHYGYYECSMQAMKNDGVVSSFFTYTGPS---DDNPWDEIDIEILGKNT 690

Qy     67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
            | |  |    |     || : :   : :|:|||| :| ||||: | |||:|| :
Db    691 TQVQFNYYTNGQGKH---EKLYDLGFDSSEAYHTYGFDWQPNYIAWYVDGREVYRA---- 743

Qy    127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
          ::  | | :  | |   :   |:  |:  : ||     ||   |
Db    744 TQDIPKTPGKIMMNAWPGLTVDDWLKAFN-GRTPLTAHYQWVTYNK 788


RESULT 6
Q93GE7
ID   Q93GE7      PRELIMINARY;      PRT;   214 AA.
AC   Q93GE7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254959; AAK50612.2; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER      1      1
SQ   SEQUENCE   214 AA;  24175 MW;  3DB192D2F6B4CAA3 CRC64;

  Query Match             16.8%;  Score 241.5;  DB 2;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 8.4e-12;
  Matches   60;  Conservative   24;  Mismatches   74;  Indels   11;  Gaps    6;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db     55 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 111

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
            ||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
Db    112 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 167

Qy    123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
          |:    |   :  |||:      |:| ::    ||:   :||:  |
Db    168 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 213


RESULT 7
Q93GE6
ID   Q93GE6      PRELIMINARY;      PRT;   214 AA.
AC   Q93GE6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254961; AAK50614.2; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER      1      1
SQ   SEQUENCE   214 AA;  24131 MW;  1E9D8BFA4300EBA5 CRC64;

  Query Match             16.4%;  Score 235.5;  DB 2;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 2.5e-11;
  Matches   60;  Conservative   23;  Mismatches   75;  Indels   11;  Gaps    6;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db     55 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIESL 111

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
          ||:    | |  |  ||     ||   :   |  |:| |  :| || ::| |||| :: |
Db    112 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHMYAFDWQPNSIKWYVDGQ-LKHT 167

Qy    123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
            ||    |   :  |||:      |:| ::    ||:   :||:  |
Db    168 ATTQVPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 213


RESULT 8
Q45691
ID   Q45691      PRELIMINARY;      PRT;   239 AA.
```

```
          |  :   || | :  ||| ||  || |   ||||| | |  |    || || |:||| ||
Db      99 APPYKAGELRTNDFYQYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFLG 157

Qy      64 KNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
           |:    | |  |    |    :| :: :   | ::|:|| ||  : :  |:|: |
Db     158 KDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA- 213

Qy     124 GGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP----GQGE 177
                |:  | | :  |||       | | ||    ||     ||   :|||     : |
Db     214 ---TENIPQTPQKIMMNLWPGVGVDEWTGVFDGDNTPLHADYEWV---RYTPLEALDEEE 267

Qy     178 GGSDFTLDWT----------------------DNFDTFDGSRWG-KGDWTFDGNRVDLT 213
           | |   :                        : |:||:   |     || :|   : |
Db     268 SGDDNEEEPVEEVEEEPAEDEEVSVRPIGSAIYETFNTFNEDIWSIAHGWT-NGQMFNAT 326

Qy     214 --DKNIYSRDGMLILALTRKGQESFNGQVPRDDEPA 247
             |  |   :|::  |: ::           ||| |
Db     327 WYDSQITFSNGIMRFAIDKE----------DDEEA 351


RESULT 2
Q9APD8
ID   Q9APD8      PRELIMINARY;      PRT;   256 AA.
AC   Q9APD8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-beta-1,3-1,4-glucanase.
GN   BGLBC1.
OS   Bacillus circulans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1397;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC21367;
RA   Lee D.-S., Kim J.-Y., Kim H.-B.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF306531; AAG53947.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
SQ   SEQUENCE   256 AA;  27604 MW;  989E50C09F6E4531 CRC64;

  Query Match             17.3%;  Score 248.5;  DB 2;  Length 256;
  Best Local Similarity   34.7%;  Pred. No. 2.9e-12;
  Matches   61;  Conservative   28;  Mismatches   70;  Indels   17;  Gaps    7;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           | | :: ||  | ::  ||: |||:| |  :| |:|:| |   |  | |    |:|||:|
Db      91 SGKPYASAEYATTQKYGYGRVEARIKAAKGTGLVTSLFTY---SGAAPGTSNDEIDIEIL 147

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA--FHTYGLEWTPNYVRWTVDGQEVR 120
           ||:    ::|  |   |   |      :    | :  || |  ||:|  ::| |||: |
Db     148 GKDTTKMETNYFTNGVGGHST-----VIDLGFDASLDFHDYAFEWSPTSIKWYVDGRLVH 202

Qy     121 KTEGGQVSNLTGTQG-LRFNLWSSESAA--WVGQFDESKLPLFQFINWVKVYKYTP 173
             || |    |  : | :  ||||    |  | |:|     |:  : :|:   |:||
Db     203 -TETGSRGPLPTSPGYIMVNLWSGAGPAEIWTGKFTYPGHPIRAYYDWI---KFTP 254


RESULT 3
Q45648
ID   Q45648      PRELIMINARY;      PRT;   276 AA.
AC   Q45648;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Lichenase precursor.
GN   BGA1.
OS   Bacillus sp.
```

```
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1409;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=137;
RX   MEDLINE=94288605; PubMed=7517127;
RA   Tabernero C., Coll P.M., Fernandez-Abalos J.M., Perez P.,
RA   Santamaria R.I.;
RT   "Cloning and DNA sequencing of bgaA, a gene encoding an endo-beta-1,3-
RT   1, 4-glucanase, from an alkalophilic Bacillus strain (N137).";
RL   Appl. Environ. Microbiol. 60:1213-1220(1994).
DR   EMBL; Z12151; CAA78135.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Signal.
FT   SIGNAL        1     31       POTENTIAL.
FT   CHAIN        32    276       LICHENASE.
SQ   SEQUENCE   276 AA;  31770 MW;  D047F8A34CA9EBE2 CRC64;

  Query Match             17.2%;  Score 247.5;  DB 2;  Length 276;
  Best Local Similarity   36.1%;  Pred. No. 3.9e-12;
  Matches   61;  Conservative   20;  Mismatches   75;  Indels   13;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           :   || | :   || ||  || | ::|||| | |   |    :  || |:||| |||:
Db      96 YKAGELRTNDFYHYGLFEVSMKPAKSTGTVSSFFTY-TGPWDWENDPWDEIDIEFLGKDT 154

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |    :| :| :   |   |:||  || |  :|| |:|: |
Db     155 TKIQFNYFTNGVGG---NEHYHELGFDAADDFNTYAFEWRPESIRWFVNGELVHTA---- 207

Qy     127 VSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             |:  | | :  |||       | |:|:    |:    :||   ||||
Db     208 TENIPQTPQKIMMNLWPGIGVDGWTGRFNGEDTPVVTQYDWV---KYTP 253


RESULT 4
Q93GE8
ID   Q93GE8      PRELIMINARY;      PRT;   214 AA.
AC   Q93GE8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254957; AAK50610.2; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   214 AA;  24159 MW;  3DB186D3B9D991CF CRC64;

  Query Match             16.9%;  Score 242.5;  DB 2;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 7e-12;
  Matches   60;  Conservative   24;  Mismatches   74;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   | :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      55 SYNKFDCGENRSVQTYGYGPYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 111

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:13:37 ; Search time 23.7951 Seconds
(without alignments)
2312.009 Million cell updates/sec

Title: US-09-654-652A-2
Perfect score: 1439
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....PNSSSVDKLAAALEHHHHHH 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	266.5	18.5	851	16	Q9K7X5	Q9k7x5 bacillus ha
2	248.5	17.3	256	2	Q9APD8	Q9apd8 bacillus ci
3	247.5	17.2	276	2	Q45648	Q45648 bacillus sp
4	242.5	16.9	214	2	Q93GE8	Q93ge8 uncultured
5	242.5	16.9	802	2	Q9S310	Q9s310 ruminococcu
6	241.5	16.8	214	2	Q93GE7	Q93ge7 uncultured
7	235.5	16.4	214	2	Q93GE6	Q93ge6 uncultured
8	234.5	16.3	239	2	Q45691	Q45691 bacillus su
9	232.5	16.2	242	2	Q8RMP0	Q8rmp0 bacillus su
10	230	16.0	212	2	Q9FDC9	Q9fdc9 paenibacill
11	230	16.0	237	2	O07856	O07856 streptococc
12	221	15.4	246	16	Q97FD3	Q97fd3 clostridium
13	219	15.2	205	2	Q93UI1	Q93ui1 uncultured
14	206	14.3	289	16	Q8U8N5	Q8u8n5 agrobacteri
15	200	13.9	293	16	Q98C78	Q98c78 rhizobium l
16	190.5	13.2	302	16	Q9K7X6	Q9k7x6 bacillus ha
17	189	13.1	245	3	O14412	O14412 orpinomyces
18	184	12.8	282	10	Q9FI31	Q9fi31 arabidopsis
19	172	12.0	282	10	O80803	O80803 arabidopsis
20	171.5	11.9	277	10	Q9M0D1	Q9m0d1 arabidopsis
21	169.5	11.8	282	10	Q9M0D2	Q9m0d2 arabidopsis
22	169	11.7	642	16	Q9WXN1	Q9wxn1 thermotoga
23	168.5	11.7	277	10	Q94A49	Q94a49 arabidopsis
24	167.5	11.6	302	3	O42800	O42800 aspergillus
25	167	11.6	646	2	Q60039	Q60039 thermotoga
26	163	11.3	286	10	Q38910	Q38910 arabidopsis
27	161	11.2	284	10	Q9FKL8	Q9fkl8 arabidopsis
28	159	11.0	286	10	Q8W4M6	Q8w4m6 arabidopsis
29	159	11.0	287	10	Q9ZSU4	Q9zsu4 arabidopsis
30	157.5	10.9	92	3	Q9HGU1	Q9hgu1 aspergillus
31	157.5	10.9	94	3	Q9P420	Q9p420 trichoderma
32	156	10.8	292	10	Q9ZRV1	Q9zrv1 fagus sylva
33	154.5	10.7	277	10	Q38907	Q38907 arabidopsis
34	154.5	10.7	284	10	Q9SEB0	Q9seb0 arabidopsis
35	154.5	10.7	306	2	Q51333	Q51333 oerskovia x
36	154	10.7	845	2	Q9KWF3	Q9kwf3 clostridium
37	154	10.7	845	16	Q8XNF8	Q8xnf8 clostridium
38	150.5	10.5	269	10	Q39148	Q39148 arabidopsis
39	149	10.4	163	2	Q93UI2	Q93ui2 uncultured
40	146.5	10.2	284	10	Q38857	Q38857 arabidopsis
41	145.5	10.1	285	10	Q9FKL9	Q9fkl9 arabidopsis
42	143	9.9	435	2	O68641	O68641 oerskovia x
43	142	9.9	422	3	Q05790	Q05790 saccharomyc
44	141.5	9.8	285	10	Q8S902	Q8s902 glycine max
45	141.5	9.8	289	10	P93669	P93669 hordeum vul

ALIGNMENTS

RESULT 1
Q9K7X5
ID Q9K7X5 PRELIMINARY; PRT; 851 AA.
AC Q9K7X5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Endo-beta-1,3-1,4 glucanase (licheninase) (EC 3.2.1.73).
GN BGLS OR BH3232.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001518; BAB06951.1; -.
DR HSSP; P23904; 1AJK.
DR InterPro; IPR000757; Glyco_hydro_16.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00722; Glyco_hydro_16; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR PRINTS; PR00737; GLHYDRLASE16.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 3.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 851 AA; 98142 MW; C1C4F0F158400285 CRC64;

Query Match 18.5%; Score 266.5; DB 16; Length 851;
Best Local Similarity 29.7%; Pred. No. 5.1e-13;
Matches 82; Conservative 31; Mismatches 108; Indels 55; Gaps 12;

Qy 4 AKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLG 63

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 9.5%; Score 172.5; DB 21; Length 277;
Best Local Similarity 26.6%; Pred. No. 2.5e-06;
Matches 54; Conservative 34; Mismatches 74; Indels 41; Gaps 9;

```
Qy     24 NVSAKDFSGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVD 81
          ::|    ||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:|
Db     48 SLSLDKSSGSGFQSNQEFLYGKAEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEID 101

Qy     82 IEVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVD 138
           | ||    :| :  :|: | |    |  : |    | |:  |||| : | |  : :|||
Db    102 FEFLGNISGHPYTLHTNVYT-KGSGDKEQQFHLWFDPTAN--FHTYCITWNPQRIIFTVD 158

Qy    139 GQEVRKTEGGQVSNLTGTQGLRF----------NLWSSESAAWVG---QFDESKLPLFQF 185
          |  :|:    :      ::|: |          :|| :|  |  |   : | || |   :
Db    159 GIPIREFMNAE------SRGVPFPTKQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAY 212

Qy    186 IN--------WVKVYKYTPGQGE 200
                    ||      |   :
Db    213 YRNYNVEGCVWVNGKSVCPANSQ 235
```

Search completed: January 9, 2003, 12:16:20
Job time : 38.3542 secs

```
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   282 AA;

  Query Match             9.5%;  Score 173;  DB 23;  Length 282;
  Best Local Similarity   25.6%;  Pred. No. 2.3e-06;
  Matches   64;  Conservative   36;  Mismatches   92;  Indels   58;  Gaps   12;

Qy      24 NVSAKDFSGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVD 81
           ::|    ||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:|
Db      53 SLSLDKSSGSGFQSNQEFLYGKAEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEID 106

Qy      82 IEVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVD 138
           | ||    :| :  :|: | |    |  : |    |  :  ||||  : | |  : :|||
Db     107 FEFLGNISGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVN--FHTYCITWNPQRIIFTVD 163

Qy     139 GQEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKV 191
           |  :|: :  :   :     | :|   :|| :|  |  |   : | || |   |
Db     164 GIPIREFKNPEAIGVPFPTRQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAF------ 217

Qy     192 YKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILAL 251
           |:                  |::  ||  |  |  :   |    | |            |
Db     218 YR-----------------NYNV-DGCVWANGKSSCSANSPWFTQK-----------L 246

Qy     252 TRKGQESFNG 261
             ||    |
Db     247 DSNGQTRMKG 256


RESULT 15
AAG52114
ID   AAG52114 standard; Protein; 277 AA.
XX
AC   AAG52114;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 66212.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 9.5%; Score 173; DB 21; Length 282;
Best Local Similarity 25.6%; Pred. No. 2.3e-06;
Matches 64; Conservative 36; Mismatches 92; Indels 58; Gaps 12;

```
Qy     24 NVSAKDFSGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVD 81
          ::|    ||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:|
Db     53 SLSLDKSSGSGFQSNQEFLYGKAEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEID 106

Qy     82 IEVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVD 138
           | ||    :| :  :|: | |    |  : |    |  :  |||| : | |  : :|||
Db    107 FEFLGNISGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVN--FHTYCITWNPQRIIFTVD 163

Qy    139 GQEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKV 191
          |  :|: :  :   :     | :|   :|| :|  |  |   : | || |   |        .
Db    164 GIPIREFKNPEAIGVPFPTRQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAF------ 217

Qy    192 YKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILAL 251
          |:                  |::  ||  |  |  :   |    | |            |
Db    218 YR-----------------NYNV-DGCVWANGKSSCSANSPWFTQK-----------L 246

Qy    252 TRKGQESFNG 261
             ||    |
Db    247 DSNGQTRMKG 256
```

RESULT 14
ABB91456
ID ABB91456 standard; Protein; 282 AA.
XX
AC ABB91456;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 667.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 667; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant

```
SQ   Sequence   282 AA;

  Query Match             10.1%;  Score 185;  DB 23;  Length 282;
  Best Local Similarity   27.4%;  Pred. No. 2.3e-07;
  Matches   59;  Conservative   34;  Mismatches   98;  Indels   24;  Gaps    9;

Qy      24 NVSAKDFSGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVD 81
           ::|   |||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:|
Db      53 SLSLDKFSGSGFQSHQEFLYGKVEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEID 106

Qy      82 IEVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVD 138
             | ||    :| :  :|: | |    |  : |    |  |  |||| : | |  | :|:|
Db     107 FEFLGNISGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVD--FHTYCIIWNPQRVIFTID 163

Qy     139 GQEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKV 191
           |  :|: :  :   :     | :|   :|| :|  |  |   : | || |   |     |
Db     164 GIPIREFKNSEALGVPFPKHQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAFYRNYNV 223

Qy     192 YKYTPGQGEGGSDFTLDWTDNFDTFDGS---RWGK 223
                 |:        |      | |    :| :
Db     224 DACVWSNGKSSCSANSSWFTQVLDFKGKNRVKWAQ 258


RESULT 13
AAG32464
ID   AAG32464 standard; Protein; 282 AA.
XX
AC   AAG32464;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 39167.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.1%; Score 185; DB 21; Length 282;
Best Local Similarity 27.4%; Pred. No. 2.3e-07;
Matches 59; Conservative 34; Mismatches 98; Indels 24; Gaps 9;

```
Qy     24 NVSAKDFSGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVD 81
          ::|   |||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:|
Db     53 SLSLDKFSGSGFQSHQEFLYGKVEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEID 106

Qy     82 IEVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVD 138
           | ||    :| :  :|: | |    |  : |    |  |  |||| : | |  | :|:|
Db    107 FEFLGNISGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVD--FHTYCIIWNPQRVIFTID 163

Qy    139 GQEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKV 191
          |  :|: :  :   :    | :|   :|| :|  |  |   : | || |   |     |
Db    164 GIPIREFKNSEALGVPFPKHQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAFYRNYNV 223

Qy    192 YKYTPGQGEGGSDFTLDWTDNFDTFDGS---RWGK 223
                |:       |      | |    :| :
Db    224 DACVWSNGKSSCSANSSWFTQVLDFKGKNRVKWAQ 258
```

RESULT 12
ABB93737
ID ABB93737 standard; Protein; 282 AA.
XX
AC ABB93737;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2948.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 2948; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX

```
CC   of grain (e.g. barley or wheat) in the brewing and fermentation
CC   industries to increase carbon substrate availability and to maximise
CC   production of desired products.
XX
SQ   Sequence   245 AA;

  Query Match             10.4%;  Score 190.5;  DB 19;  Length 245;
  Best Local Similarity   31.6%;  Pred. No. 6.5e-08;
  Matches   54;  Conservative   17;  Mismatches   87;  Indels   13;  Gaps    6;

Qy      21 LTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEV 80
           || :     ::  |  |     || |:  ||     | ||| | |   |   ||  | |:
Db      79 LTIDRDGSGYTCGEYRTKNYYGYGMFQVNMKPIKNPGVVSSFFTYTGPS---DGTKWDEI 135

Qy      81 DIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQ 140
           ||| || :    |  |  |    |    |  | :   | | |||||  |  | : | |||
Db     136 DIEFLGYDTTKVQFNYYTNGQGHH---EHIHYLGFDASQGFHTYGFFWARNSITWYVDGT 192

Qy     141 EVRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWV 189
           |         |:  | | :  | |:      |:  |:  :  :  : :||
Db     193 AVYTA----YDNIPDTPGKIMMNAWNGIGVDDWLRPFN-GRTNISAYYDWV 238


RESULT 11
AAG48467
ID   AAG48467 standard; Protein; 282 AA.
XX
AC   AAG48467;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 61207.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
```

```
CC   protein accumulates in the ripe (but ungerminated) seeds.  The
CC   proportion of these to malted seeds is increased therefore, without
CC   an unacceptable increase in viscosity.
XX
SQ   Sequence   237 AA;

  Query Match             12.5%;  Score 228;  DB 11;  Length 237;
  Best Local Similarity   34.3%;  Pred. No. 4.4e-11;
  Matches   60;  Conservative   25;  Mismatches   74;  Indels   16;  Gaps    8;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: :   |  ||  :     || :|  || |  :| ||| | |      | |  | |:||
Db      75 TSSAYNKFDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDI 131

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |   |     ||  ::   | : ||||  :| | |::| |||  :
Db     132 EFLGKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VL 187

Qy     143 RKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           : |     :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     188 KHT---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 235


RESULT 9
AAR06622
ID   AAR06622 standard; protein; 237 AA.
XX
AC   AAR06622;
XX
DT   09-JAN-1991  (first entry)
XX
DE   Hybrid (1,3-1,4)-pre-beta-glucanase.
XX
KW   Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.
XX
OS   Bacillus amyloliquefaciens, Bacillus macerans.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..127
FT                   /label=amino-terminal half of B.macerans beta-glucanase
FT   Domain          131..234
FT                   /label=carboxyl-terminal half of B.amyloliquefaciens
FT                         beta-glucanase
XX
PN   WO9009436-A.
XX
PD   23-AUG-1990.
XX
PF   16-FEB-1990;   90WO-DK00044.
XX
PR   04-AUG-1989;   89DK-0003848.
PR   16-FEB-1989;   89DD-0325800.
XX
PA   (CARL-) CARLSBERG A/S.
PA   (DEAK ) AKAD WISSENSCHAFT DDR.
XX
PI   Borriss R,  Hofemeister J,  Thomsen KK,  Olsen O,  Vonwettstein D;
XX
DR   WPI; 1990-275129/36.
DR   N-PSDB; AAQ05833.
XX
PT   New thermostable (1,3-1,4)-beta-glucanase - prepd. using hybrid
PT   gene obtd. using Bacillus amyloliquefaciens and B.macerans genes
XX
PS   Disclosure; page 28; 84pp; English.
XX
CC   This hybrid protein is encoded by the beta-glucanase-H2 gene.
CC   Following processing of the signal peptide the mature protein
CC   is produced, comprising the amino terminus of the B.macerans
CC   beta-glucanase and the carboxyl-terminal half of the B.amylo-
CC   liquefaciens beta-glucanase.  This hybrid protein is thermo-
CC   stable and hydrolyses beta-glycosidic linkages in (1,3-1,4)-
CC   beta-glucans.  Reducing sugars are obtd. at high temps. and
CC   thus this enzyme can be used in the mfr. of food prods., esp.
CC   beer and animal feed (e.g.for feeding poultry).  See also
CC   AAQ05833.
XX
SQ   Sequence   237 AA;

  Query Match             12.2%;  Score 222.5;  DB 11;  Length 237;
  Best Local Similarity   33.1%;  Pred. No. 1.3e-10;
  Matches   57;  Conservative   25;  Mismatches   79;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: :   |  ||  :     || :|  || |  :| ||| | |      | |  | |:||
Db      75 TSSAYNKFDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDI 131

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |   ||     ||   :   |  |:|||  :| || ::| |||| :
Db     132 EFLGKDTTKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 187

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           : |   |:    |   :  |||:      |:| ::    |::   :|::  |
Db     188 KHTATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYHWMRYRK 236


RESULT 10
AAW37884
ID   AAW37884 standard; Protein; 245 AA.
XX
AC   AAW37884;
XX
DT   20-AUG-1998  (first entry)
XX
DE   Lichenase protein.
XX
KW   Lichenase; licA; fungus; enzyme; beta-1,4-glucan bond hydrolysation;
KW   beta-1,3-linked glucan bond; grain-containing feed; grain treatment;
KW   nutrient availability; brewing; fermentation.
XX
OS   Orpinomyces sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..29
FT                   /note= "signal peptide"
FT   Protein         30..246
FT                   /note= "mature lichenase"
XX
PN   WO9814595-A1.
XX
PD   09-APR-1998.
XX
PF   03-OCT-1997;   97WO-US17811.
XX
PR   04-OCT-1996;   96US-0027882.
XX
PA   (UYGE-) UNIV GEORGIA RES FOUND INC.
XX
PI   Chen H,  Li X,  Ljungdahl LG;
XX
DR   WPI; 1998-240094/21.
DR   N-PSDB; AAV29067.
XX
PT   New isolated lichenase protein - is obtained from Orpinomyces PC-2,
PT   used for treatment of grain to improve feeds or to improve brewing
PT   and fermentation processes
XX
PS   Claim 1; Page 24-25; 41pp; English.
XX
CC   This sequence is the lichenase (licA) of Orpinomyces sp. strain PC-2
CC   of the invention. The protein was purified from a fungus or a fungal
CC   culture or from a recombinant DNA molecule having a fungal lichenase
CC   coding sequence. The lichenase enzyme hydrolyses the beta-1,4-glucan
CC   bonds adjacent to beta-1,3-linked glucan bonds, but does not cleave
CC   beta-1,4-linked glucans. They can be used for the treatment of animal
CC   grain-containing feeds to improve nutrient availability and for treatment
```

KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.
XX
OS Unidentified.
XX
PN WO200212511-A1.
XX
PD 14-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-ES00303.
XX
PR 28-JUL-2000; 2000ES-0001922.
XX
PA (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX
PI Rodriguez Munoz V, Perez Mellado R;
XX
DR WPI; 2002-217195/27.
DR N-PSDB; ABL53375.
XX
PT Producing nucleic acid encoding lichenase, useful for producing enzymes
PT for improving filtration in brewing, comprises isolating the nucleic
PT acid from soil bacteria -
XX
PS Claim 18; Page 21-22; 27pp; Spanish.
XX
CC The present sequence is a protein sequence from a soil bacterium for an
CC enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)
CC activity. The enzyme is useful in brewing and for degrading
CC beta-glucanases that cause problems during filtration.
XX
SQ Sequence 214 AA;

Query Match 12.6%; Score 229; DB 23; Length 214;
Best Local Similarity 36.1%; Pred. No. 3.2e-11;
Matches 56; Conservative 22; Mismatches 67; Indels 10; Gaps 5;

```
Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db      60 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 116

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :
Db     117 EFLGKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 172

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 176
           : |   |:    |  :  |||:      |:| ::
Db     173 KHTATTQIPAAPGK--IMMNLWNGTGVDDWLGSYN 205
```

RESULT 7
AAR03775
ID AAR03775 standard; protein; 234 AA.
XX
AC AAR03775;
XX
DT 31-JUL-1990 (first entry)
XX
DE Thermostable beta-glucanase.
XX
KW Thermostable beta-glucanase; ss; Bacillus subtilis.
XX
OS Bacillus macerans.
XX
PN DD272102-A.
XX
PD 27-SEP-1989.
XX
PF 12-MAY-1988; 88DD-0315706.
XX
PR 12-MAY-1988; 88DD-0315706.
XX
PA (DEAK) AKAD WISSENSCHAFT DDR.
XX
PI Borriss R;
XX
DR WPI; 1990-067913/10.
DR N-PSDB; AAQ03519.
XX
PT Thermostable beta-glucanase production -
PT using Bacillus subtilis transformed with gene from Bacillus
PT macerans.
XX
PS Fig 1; ; 9pp; German.
XX
CC The gene encoding Bacillus macerans thermostable
CC beta-glucanase is expressed in Bacillus subtilis. The
CC enzyme is useful for lowering the viscosity of brewing
CC mashes and in the production of feedstuff.
XX
SQ Sequence 234 AA;

Query Match 12.5%; Score 228; DB 11; Length 234;
Best Local Similarity 34.3%; Pred. No. 4.3e-11;
Matches 60; Conservative 25; Mismatches 74; Indels 16; Gaps 8;

```
Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: :   | ||  :     || :|  || |  :| ||| | |      | |  | |:||
Db      72 TSSAYNKFDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDI 128

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |    |     ||  ::   | : ||||  :| | |::| |||  :
Db     129 EFLGKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VL 184

Qy     143 RKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
           : |     :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db     185 KHT---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 232
```

RESULT 8
AAR05803
ID AAR05803 standard; protein; 237 AA.
XX
AC AAR05803;
XX
DT 08-NOV-1990 (first entry)
XX
DE Heat-stable endo-beta-1,3-1,4-glucanase.
XX
KW Endo-beta-1,3-1,4-glucanase; barley; brewing.
XX
OS Bacillus macerans.
XX
PN DD275704-A.
XX
PD 31-JAN-1990.
XX
PF 23-SEP-1988; 88DD-0320082.
XX
PR 23-SEP-1988; 88DD-0320082.
XX
PA (DEAK) AKAD WISSENSCHAFT DDR.
XX
PI Borriss R, Wobus U, Mendel R-R, Baumlein H;
XX
DR WPI; 1990-210631/28.
DR N-PSDB; AAQ05167.
XX
PT Prepn. of barley plants expressing heat stable beta-glucanase -
PT by transforming cells with appropriate vector then regeneration
PT giving seeds useful in brewing without conversion to malt.
XX
PS Disclosure; ; p; German.
XX
CC The corresp. gene and the plant promoter region are inserted into an
CC expression vector and used to transform barley cells. The transformants
CC are used to regenerate barley plants which are useful in brewing. This

```
Qy    143 RKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
          : |     :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db    165 KHT---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 212
```

RESULT 4
AAR06621
ID AAR06621 standard; protein; 239 AA.
XX
AC AAR06621;
XX
DT 09-JAN-1991 (first entry)
XX
DE Hybrid (1,3-1,4)-pre-beta-glucanase.
XX
KW Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.
XX
OS Bacillus amyloliquefaciens, Bacillus macerans.
XX
FH Key Location/Qualifiers
FT Domain 1..129
FT /label=amino terminal of beta-amyloliquefaciens
FT Domain 133..236
FT /label=carboxyl-terminal of B.macerans
XX
PN WO9009436-A.
XX
PD 23-AUG-1990.
XX
PF 16-FEB-1990; 90WO-DK00044.
XX
PR 04-AUG-1989; 89DK-0003848.
PR 16-FEB-1989; 89DD-0325800.
XX
PA (CARL-) CARLSBERG A/S.
PA (DEAK) AKAD WISSENSCHAFT DDR.
XX
PI Borriss R, Hofemeister J, Thomsen KK, Olsen O, Vonwettstein D;
XX
DR WPI; 1990-275129/36.
DR N-PSDB; AAQ05832.
XX
PT New thermostable (1,3-1,4)-beta-glucanase - prepd. using hybrid
PT gene obtd. using Bacillus amyloliquefaciens and B.macerans genes
XX
PS Disclosure; page 26; 84pp; English.
XX
CC This hybrid protein is encoded by the beta-glucanase-H1 gene.
CC Following processing of the signal peptide the mature protein
CC is produced, comprising the amino terminus of the amylolique-
CC faciens beta-glucanase and the carboxyl-terminal half of the
CC B.macerans beta-glucanase. This hybrid protein is thermostable
CC and hydrolyses beta-glycosidic linkages in (1,3-1,4)-beta-glucans.
CC Reducing sugars are obtd. at high temps. and thus this enzyme can
CC be used in the mfr. of food prods., esp. beer and animal feed (eg
CC for feeding poultry). See also AAQ05833.
XX
SQ Sequence 239 AA;

Query Match 13.6%; Score 248; DB 11; Length 239;
Best Local Similarity 35.4%; Pred. No. 9.3e-13;
Matches 62; Conservative 27; Mismatches 70; Indels 16; Gaps 8;

```
Qy     23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
          |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db     77 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 133

Qy     83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
          | |||:    | |  |  |    ||  ::   | : ||||  :| | |::| |||  :
Db    134 EFLGKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VL 189

Qy    143 RKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 195
```

```
          : |     :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db    190 KHT---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 237
```

RESULT 5
ABB76858
ID ABB76858 standard; Protein; 208 AA.
XX
AC ABB76858;
XX
DT 27-JUN-2002 (first entry)
XX
DE Bacterial lichenase #1.
XX
KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.
XX
OS Unidentified.
XX
PN WO200212511-A1.
XX
PD 14-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-ES00303.
XX
PR 28-JUL-2000; 2000ES-0001922.
XX
PA (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX
PI Rodriguez Munoz V, Perez Mellado R;
XX
DR WPI; 2002-217195/27.
DR N-PSDB; ABL53374.
XX
PT Producing nucleic acid encoding lichenase, useful for producing enzymes
PT for improving filtration in brewing, comprises isolating the nucleic
PT acid from soil bacteria -
XX
PS Claim 18; Page 18-19; 27pp; Spanish.
XX
CC The present sequence is a protein sequence from a soil bacterium for an
CC enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)
CC activity. The enzyme is useful in brewing and for degrading
CC beta-glucanases that cause problems during filtration.
XX
SQ Sequence 208 AA;

Query Match 12.6%; Score 229; DB 23; Length 208;
Best Local Similarity 36.1%; Pred. No. 3e-11;
Matches 56; Conservative 22; Mismatches 67; Indels 10; Gaps 5;

```
Qy     23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
          |: |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||
Db     55 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 111

Qy     83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
          | |||:    | |  |  |    ||  ::   | : ||||  :| | |::| |||  :
Db    112 EFLGKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 167

Qy    143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 176
          : |   |:    |  :  |||:      |:| ::
Db    168 KHTATTQIPAAPGK--IMMNLWNGTGVDDWLGSYN 200
```

RESULT 6
ABB76859
ID ABB76859 standard; Protein; 214 AA.
XX
AC ABB76859;
XX
DT 27-JUN-2002 (first entry)
XX
DE Bacterial lichenase #2.
XX

```
XX
PS   Claim 6; Page 5; 8pp; German.
XX
CC   This sequence represents a novel beta-glucanase isolated from Bacillus
CC   alkalophilus DSM 9956. This enzyme is useful for removing glucan and/or
CC   lichenan from membranes and equipment in the food industry, especially
CC   the brewing industry.
XX
SQ   Sequence   308 AA;

  Query Match            14.1%;  Score 257.5;  DB 20;  Length 308;
  Best Local Similarity  36.5%;  Pred. No. 2.1e-13;
  Matches   62;  Conservative   22;  Mismatches   73;  Indels   13;  Gaps    6;

Qy      29 DFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN 88
           ::   || | :  ||| ||  || | ::|||||:| |  |    |  || |:||| |||:
Db      96 EYKAGELRTNQFYQYGLFEVNMKPAKSTGTVSSLFTY-TGPWDWDNDPWDEIDIEFLGKD 154

Qy      89 PGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGG 148
             | |  |   |    :| :| :   | ::|:||  || |  : | |:|: |
Db     155 TTRVQFNYFTNGVG---NNEHYHELGFDASESFNTYAFEWRPESISWYVNGELVYTA--- 208

Qy     149 QVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 196
              |:  | | :  |||       | | ||    |:    :||   :||
Db     209 -TENIPQTPQKIMMNLWPGIGVDGWTGVFDGEDTPVVTEYDWV---RYTP 254


RESULT 2
AAP95000
ID   AAP95000 standard; Protein; 242 AA.
XX
AC   AAP95000;
XX
DT   12-FEB-1997  (first entry)
XX
DE   Bacillus subtilis lichenase.
XX
KW   Beer production; fermentation; barley; beta-glucan; hydrolysis;
KW   lichenase.
XX
OS   Bacillus subtilis Y-25.
XX
PN   JP01067181-A.
XX
PD   13-MAR-1989.
XX
PF   08-SEP-1987;   87JP-0224615.
XX
PR   08-SEP-1987;   87JP-0224615.
XX
PA   (ASAK ) ASAHI BREWERIES KK.
XX
DR   WPI; 1989-119863/16.
DR   N-PSDB; AAN95000.
XX
PT   Recombinant plasmid used in beer prodn. - obtd. by integrating
PT   lichenase gene derived from Bacillus subtilis, into vector
XX
PS   Disclosure; Fig 4; 7pp; Japanese.
XX
CC   The lichenase gene from Bacillus subtilis Y-25 is used for
CC   transforming Bacillus hosts so that they show increased lichenase
CC   expression. The recombinant lichenase enzyme produced by the
CC   transformants is useful in beer production for decomposing beta-
CC   glucan from barley.
XX
SQ   Sequence   242 AA;

  Query Match            13.7%;  Score 250.5;  DB 10;  Length 242;
  Best Local Similarity  35.5%;  Pred. No. 5.8e-13;
  Matches   61;  Conservative   25;  Mismatches   75;  Indels   11;  Gaps    6;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||
Db      80 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDI 136

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :
Db     137 EFLGKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-L 192

Qy     143 RKTEGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 193
           : |   |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     193 KHTATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 3
AAE07317
ID   AAE07317 standard; Protein; 214 AA.
XX
AC   AAE07317;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Barley recombinant, thermostable (1,3-1,4)-beta-glucanase enzyme.
XX
KW   Barley; transgenic barley malt; carbohydrate degrading enzyme; chicken;
KW   (1,3-1,4)-beta-glucanase; nutritional value; foodstuff; thermostable.
XX
OS   Hordeum vulgare.
XX
PN   WO200159141-A2.
XX
PD   16-AUG-2001.
XX
PF   09-FEB-2001; 2001WO-US04222.
XX
PR   10-FEB-2000; 2000US-0181473.
PR   09-NOV-2000; 2000US-0247126.
XX
PA   (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Von Wettstein D,  Huang J,  Horvath H;
XX
DR   WPI; 2001-497082/54.
XX
PT   New barley based foodstuff for animals, i.e. chicken comprising
PT   recombinant carbohydrate degrading enzyme (1,3-1,4)-beta-glucanase -
XX
PS   Claim 8; Page 37-38; 43pp; English.
XX
CC   The present invention relates to a food stuff comprising barley feed and
CC   transgenic barley malt where the transgenic barley malt comprises a
CC   recombinant carbohydrate degrading enzyme comprising a (1,3-1,4)-beta-
CC   glucanase. The food stuff is useful to feed animal, preferably chickens
CC   Barley is cheaper than corn, the principal foodstuff of chickens.
CC   Chickens cannot efficiently utilise barley as an energy source as they
CC   do not possess a gut enzyme that depolymerises beta-D-glucan, major
CC   carbohydrate present in the barley endosperm. The invention provides a
CC   barley based foodstuff which comprises a recombinant carbohydrate
CC   degrading enzyme improving the nutritional value of the foodstuff.
CC   The present sequence is barley recombinant, thermostable
CC   (1,3-1,4)-beta-glucanase enzyme.
XX
SQ   Sequence   214 AA;

  Query Match            13.6%;  Score 248;  DB 22;  Length 214;
  Best Local Similarity  35.4%;  Pred. No. 8e-13;
  Matches   62;  Conservative   27;  Mismatches   70;  Indels   16;  Gaps    8;

Qy      23 TNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDI 82
           |: |   |   |  :::   || :| ||| |  :| ||| | |   :   | || |:||
Db      52 TSPSYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDI 108

Qy      83 EVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEV 142
           | |||:    | |  |  |     ||  ::   | : ||||  :| | |::| |||  :
Db     109 EFLGKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VL 164
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:57 ; Search time 36.3542 Seconds
(without alignments)
1279.206 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKKTAVKSALAVAAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	257.5	14.1	308	20	AAW93001	B. alkalophilus be
2	250.5	13.7	242	10	AAP95000	Bacillus subtilis
3	248	13.6	214	22	AAE07317	Barley recombinant
4	248	13.6	239	11	AAR06621	Hybrid (1,3-1,4)-p
5	229	12.6	208	23	ABB76858	Bacterial lichenas
6	229	12.6	214	23	ABB76859	Bacterial lichenas
7	228	12.5	234	11	AAR03775	Thermostable beta-
8	228	12.5	237	11	AAR05803	Heat-stable endo-b
9	222.5	12.2	237	11	AAR06622	Hybrid (1,3-1,4)-p
10	190.5	10.4	245	19	AAW37884	Lichenase protein.
11	185	10.1	282	21	AAG48467	Arabidopsis thalia
12	185	10.1	282	23	ABB93737	Herbicidally activ
13	173	9.5	282	21	AAG32464	Arabidopsis thalia
14	173	9.5	282	23	ABB91456	Herbicidally activ
15	172.5	9.5	277	21	AAG52114	Arabidopsis thalia
16	172.5	9.5	277	23	ABB93169	Herbicidally activ
17	170.5	9.3	269	21	AAG18644	Arabidopsis thalia
18	170.5	9.3	282	21	AAG18643	Arabidopsis thalia
19	170.5	9.3	282	23	ABB93168	Herbicidally activ
20	170.5	9.3	298	21	AAG18642	Arabidopsis thalia
21	164	9.0	286	23	ABB93103	Herbicidally activ
22	161	8.8	284	23	ABB93853	Herbicidally activ
23	159.5	8.7	287	23	ABB93104	Herbicidally activ
24	157.5	8.6	269	21	AAG20347	Arabidopsis thalia
25	157.5	8.6	269	21	AAG20350	Arabidopsis thalia
26	157.5	8.6	269	21	AAG52653	Arabidopsis thalia
27	157.5	8.6	269	23	ABB93167	Herbicidally activ
28	157.5	8.6	290	21	AAG20346	Arabidopsis thalia
29	157.5	8.6	297	21	AAG52652	Arabidopsis thalia
30	157.5	8.6	300	21	AAG20349	Arabidopsis thalia
31	156.5	8.6	269	21	AAG20329	Arabidopsis thalia
32	156.5	8.6	299	21	AAG20328	Arabidopsis thalia
33	155	8.5	845	22	AAB98062	Clostridium perfri
34	154.5	8.5	284	23	ABB93854	Herbicidally activ
35	154.5	8.5	306	17	AAR97362	Oerskovia beta-1,3
36	154	8.4	247	18	AAW11593	Protein encoded by
37	154	8.4	247	20	AAY08308	A. thaliana merist
38	145.5	8.0	285	23	ABB93852	Herbicidally activ
39	145.5	8.0	307	21	AAG40361	Arabidopsis thalia
40	145.5	8.0	310	21	AAG40360	Arabidopsis thalia
41	145.5	8.0	310	23	ABB90942	Herbicidally activ
42	145.5	8.0	312	21	AAG40359	Arabidopsis thalia
43	145	7.9	263	18	AAW29455	Oerskovia xanthine
44	145	7.9	305	23	ABB91796	Herbicidally activ
45	144.5	7.9	307	21	AAG16385	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW93001
ID AAW93001 standard; Protein; 308 AA.
XX
AC AAW93001;
XX
DT 19-MAY-1999 (first entry)
XX
DE B. alkalophilus beta-glucanase protein.
XX
KW Beta-glucanase; enzyme; glucan; removal; lichenan; membrane;
KW equipment; food industry; brewing.
XX
OS Bacillus alkalophilus.
XX
PN DE19732751-A1.
XX
PD 04-FEB-1999.
XX
PF 30-JUL-1997; 97DE-1032751.
XX
PR 30-JUL-1997; 97DE-1032751.
XX
PA (HENK) HENKEL KGAA.
XX
PI Hillen W, Maurer K;
XX
DR WPI; 1999-122161/11.
DR N-PSDB; AAX02912.
XX
PT New Bacillus alkalophilus beta-glucanase protein and gene - useful
PT for removing glucan and/or lichenan from membranes in the brewing
PT industry

```
;   LENGTH: 279
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-3942

  Query Match             5.0%;  Score 91;  DB 9;  Length 279;
  Best Local Similarity   22.6%;  Pred. No. 2.2;
  Matches   67;  Conservative   39;  Mismatches  126;  Indels   64;  Gaps   13;

Qy       1 MNIKKTAVKSALAVAAAAAALTTNVSAKD----FSGAELYTLEEVQYGKFEARMKMAAAS 56
           ||   :|: :|    || |   |:  | :    |:|   : : :     | : :    |:
Db       1 MNRFPSALLAASVAGAALAIPATSAHAAESICTFTGDVGWNVRD----SFNSYLLGNIAN 56

Qy      57 GTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPG------SFQSNIITGKAGAQKTSEKH 110
           |   | : |: | :: ||     |  :  || |       |  |: |:   ||  |
Db      57 G---SAYKYKGGLDVRDG-----VQTDGNGKTPALTWPVDSVTSSKISTSGGAHWTG--- 105

Qy     111 HAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQG-LRFNLWSSE-- 167
           | : |  |       |   ||     :|   :          |:|: | |  :  | |
Db     106 HNLYPGDD-------LAEVPNNFTLDLDFSNI-------TVELSGSTGRLLVDYTSREYI 151

Qy     168 SAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWT 227
           :   :|:|   :      | : |    |         |  |       | ||     | :|
Db     152 NTHTLGEFQTGEQAELATITFAKAPDLTSNSVNVTGDVAL-------TADGVEVFGGFYT 204

Qy     228 FDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSS 283
                 :|  |: : :|               |  |  ::|   |:||:  ||||
Db     205 AGEELAPIT-LNLTADEGC-------------GSTPEPEDPETPSASSSDVSSSS 245


Search completed: January  9, 2003, 12:14:42
Job time : 10.8866 secs
```

```
Qy      56 SGTVSSMFLYQNGSEIA-----DGRPWVEVD-IEVLGKNPGSFQSNIITGKAGAQKTSEK 109
           :|:: |   |||   |       | |   |     |     ||  :    |  :: :
Db    1080 NGSIRSDISYQNIDAIVLPTLPKLRHWFMSDRFSEQGDIMAKFQGALNRAYAQLEQNTLL 1139

Qy     110 HHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSS--E 167
           |:               :| :   |||:|      :  ||    | : ||   |||
Db    1140 HNG-------------HFTKDAANWTVEG------DAHQVVLEDGKRVLRLPDWSSSVS 1179

Qy     168 SAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLD----------WTDNFDTFD 217
           :  ||  |   :| :          |||||    ||:           | :|  |
Db    1180 QTIEIENFDPDK--EYQLV--------FHGQGEG--TVTLEHGEETKYIETHTHHFANFT 1227

Qy     218 GSRWGKGDWTFDGNRVDLTDKNIYSRDGMLI---LALTRKGQESFNGQVPRDDEPAPQSS 274
           |:  :   ||: |:| :|   | | ||  :   :||           :| ||: :  ::
Db    1228 TSQ--RQGLTFESNKVTVT---ISSEDGEFLVDNIALV-------EAPLPTDDQNSEGNT 1275

Qy     275 SSAPASSSSV 284
           :|:  | :|:
Db    1276 ASSTNSDTSM 1285


RESULT 13
US-09-925-301-1154
; Sequence 1154, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1154
;   LENGTH: 156
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-301-1154

  Query Match             5.1%;  Score 93;  DB 10;  Length 156;
  Best Local Similarity   38.0%;  Pred. No. 0.7;
  Matches   30;  Conservative   14;  Mismatches   19;  Indels   16;  Gaps    3;

Qy     265 RDDEPAP-----QSSSSAPASSSSVPASSSSVPASSSSAFVPPSSSSATNAIHGMRTTPA 319
           |  || |     :||||: :||||  :||||  :||||:    ||||::::     ::|:
Db      69 RPPEPTPAKRKRRSSSSSSSSSSSSSSSSSSSSSSSSSSS---SSSSSSSSSSSSSSSPS 125

Qy     320 VAKEHRNLVNAKGAKVNPN 338
           ||         | :  ||
Db     126 PAK--------PGPQACPN 136


RESULT 14
US-09-738-626-6377
; Sequence 6377, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6377
;   LENGTH: 714
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-6377

  Query Match             5.1%;  Score 93;  DB 9;  Length 714;
  Best Local Similarity   21.2%;  Pred. No. 5.4;
  Matches   68;  Conservative   48;  Mismatches  136;  Indels   68;  Gaps   14;

Qy      30 FSGAELYTLEEVQYGKFEARMKMAAAS-----GTVSSMFLYQNGSEIA--DGRPWVEVDI 82
           :| | :: | |:  | |   ::  | |     |   ::::  | :  |  ||||: |:||
Db     436 YSSARVH-LPEIPAGNFRLTVRARAQSEELVDGVRPAIWMQNNTNFCADNDGRPYGELDI 494

Qy      83 -EVLGKNPGSFQSNIITGKAG---AQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVD 138
            |       :  | :  | ||     |  :     |   |  :| :|:|     : :|:|
Db     495 TEFYSSRVNTQYSAVHLGCAGNRPEMKLRQMEMEESMFGD--WHDWGVEVFDGQIVFTID 552

Qy     139 GQEV----RKTEGGQVSNLTG-TQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYK 193
           |: |    :   |  |:       :  | |   |   :||      |    :| :
Db     553 GKAVTSSGKDVFGNSVTPAAAPLRPAHFKLSEEEYREVIGQ------PWHLILNTM---- 602

Qy     194 YTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTR 253
                 : | |  :   || : |   |     :  |   ||:   ::
Db     603 ----VEQSGKDSWITAVDNNEAFPEHR-----FQIDHVAVDIESDSV------------ 640

Qy     254 KGQESFNGQVPRDDEPAPQSSSSAPAS---SSSVPASSSSVPASSSSAFVPPSSSSATNA 310
                       |   | :::  | :     |   |   | :: ||     |  |   |
Db     641 ------------DNVWPDAANEIPDNVGIEDSDDGSDLEVGSTGSSTAETVSWISLFTA 687

Qy     311 IHGMRTTPAVAKEH-RNLVN 329
           :  :  | |: :|  :||:|
Db     688 LSSLVFTLALNQEALQNLIN 707


RESULT 15
US-09-738-626-3942
; Sequence 3942, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3942
```

```
  Query Match             5.3%;  Score 96;  DB 10;  Length 821;
  Best Local Similarity   21.2%;  Pred. No. 3.6;
  Matches   87;  Conservative   49;  Mismatches  170;  Indels  104;  Gaps   18;

Qy       2 NIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSS 61
           |||   |:             |:| :   |  :|  | |                 | ||
Db     174 NIKSLLVE------------VNISVE--KGGSVYAKERVSLEN--------VTEATFSS 210

Qy      62 MFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAF 121
             || |   |  : | |: :   |   || |     |  |   |:   :      :
Db     211 -----NGGEQGGGGIYSEQDMLISDCNNVHFQGNAAGATAVKQCLDEEMIVLLTECVDSL 265

Qy     122 HTYGLEWTPNYVRWTVDG-----QEVRKTE-----------------GG-------QVS 151
             |: ||   :   :|      | : |:                  ||        ::
Db     266 SEDTLDSTPETEQTKSNGNQDGSSETKDTQVSESPESTPSPDDVLGKGGGIYTEKSLTIT 325

Qy     152 NLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTD 211
           :|||     |: :::| |  | | :  |      | ::  | : ||  ||:  |   :
Db     326 GITGTIDFVSNI-ATDSGA--GVFTKENLSCTN-TNSLQFLKNSAGQHGGGAYVTQTMSV 381

Qy     212 NFDTFDGSRWGK--GDWTFDGNRVDLTDK--------NIYSRDGMLILALTRKGQESFNG 261
             | :      |:  |  |    | |        |  |   :  : ||:   :   |
Db     382 TNTTSESITTPPLVGEVIFSEN----TAKGHGGGICTNKLSLSNLKTVTLTKNSAKESGG 437

Qy     262 -------QVPRDDEPAPQS-SSSAPASSSSVPAS--------SSSVPASSSSAFVPPSSS 305
                 :|  | |   : |||:|||:  | ||        |:: ||: |   :  : |
Db     438 AIFTDLASIPTTDTPESSTPSSSSPASTPEVVASAKINRFFASTAEPAAPS---LTEAES 494

Qy     306 SATNAIHGMRTTPAVAKEHRNLVNA---------KGAKVNPNGHKRYRVN 346
           |:      |   :     |::|          ||  :     |  |:|
Db     495 DQTDQTETSDTNSDIDVSIENILNVAINQNTSAKKGGAIYGKKAKLSRIN 544


RESULT 11
US-09-841-132-179
; Sequence 179, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 179
;   LENGTH: 1776
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-841-132-179

  Query Match             5.3%;  Score 96;  DB 10;  Length 1776;
  Best Local Similarity   21.2%;  Pred. No. 10;
  Matches   87;  Conservative   49;  Mismatches  170;  Indels  104;  Gaps   18;

Qy       2 NIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSS 61
           |||   |:             |:| :   |  :|  | |                 | ||
Db     187 NIKSLLVE------------VNISVE--KGGSVYAKERVSLEN--------VTEATFSS 223

Qy      62 MFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAF 121
             || |   |  : | |: :   |   || |     |  |   |:   :      :
Db     224 -----NGGEQGGGGIYSEQDMLISDCNNVHFQGNAAGATAVKQCLDEEMIVLLTECVDSL 278

Qy     122 HTYGLEWTPNYVRWTVDG-----QEVRKTE-----------------GG-------QVS 151
             |: ||   :   :|      | : |:                  ||        ::
Db     279 SEDTLDSTPETEQTKSNGNQDGSSETKDTQVSESPESTPSPDDVLGKGGGIYTEKSLTIT 338

Qy     152 NLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTD 211
           :|||     |: :::| |  | | :  |      | ::  | : ||  ||:  |   :
Db     339 GITGTIDFVSNI-ATDSGA--GVFTKENLSCTN-TNSLQFLKNSAGQHGGGAYVTQTMSV 394

Qy     212 NFDTFDGSRWGK--GDWTFDGNRVDLTDK--------NIYSRDGMLILALTRKGQESFNG 261
             | :      |:  |  |    | |        |  |   :  : ||:   :   |
Db     395 TNTTSESITTPPLVGEVIFSEN----TAKGHGGGICTNKLSLSNLKTVTLTKNSAKESGG 450

Qy     262 -------QVPRDDEPAPQS-SSSAPASSSSVPAS--------SSSVPASSSSAFVPPSSS 305
                 :|  | |   : |||:|||:  | ||        |:: ||: |   :  : |
Db     451 AIFTDLASIPTTDTPESSTPSSSSPASTPEVVASAKINRFFASTAEPAAPS---LTEAES 507

Qy     306 SATNAIHGMRTTPAVAKEHRNLVNA---------KGAKVNPNGHKRYRVN 346
           |:      |   :     |::|          ||  :     |  |:|
Db     508 DQTDQTETSDTNSDIDVSIENILNVAINQNTSAKKGGAIYGKKAKLSRIN 557


RESULT 12
US-09-738-363-4
; Sequence 4, Application US/09738363
; Patent No. US20010010932A1
;  GENERAL INFORMATION:
;       APPLICANT: Schnepf, Harry E.
;                  Schwab, George E.
;                  Payne, Jewel M.
;                  Narva, Kenneth E.
;                  Foncerrada, Luis
;       TITLE OF INVENTION: Nematicidal Proteins
;       NUMBER OF SEQUENCES: 40
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Jay M. Sanders
;            STREET: 2421 N.W. 41st Street, Suite A-1
;            CITY: Gainesville
;            STATE: FL
;            COUNTRY: USA
;            ZIP: 32606
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/738,363
;            FILING DATE: 15-Dec-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/076,137
;            FILING DATE: 12-MAY-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Sanders, Jay
;            REGISTRATION NUMBER: 39,355
;            REFERENCE/DOCKET NUMBER: MA-20CCCD3
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 352-375-8100
;            TELEFAX: 352-372-5800
;   INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1289 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: YES
;       ANTI-SENSE: NO
;       ORIGINAL SOURCE:
;            ORGANISM: BACILLUS THURINGIENSIS
;            INDIVIDUAL ISOLATE: PS17
;       IMMEDIATE SOURCE:
;            CLONE: E. coli NM522(pMYC 1628) NRRL B-18652
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-738-363-4

  Query Match             5.2%;  Score 95.5;  DB 10;  Length 1289;
  Best Local Similarity   23.6%;  Pred. No. 7.3;
  Matches   59;  Conservative   34;  Mismatches   92;  Indels   65;  Gaps   13;
```

```
          |:  | :|:
Db   1373 SSTNSDTSM 1381


RESULT 8
US-09-841-132-444
; Sequence 444, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 444
;   LENGTH: 1770
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis serovar D
US-09-841-132-444

  Query Match             5.3%;  Score 97;  DB 10;  Length 1770;
  Best Local Similarity   21.5%;  Pred. No. 8.4;
  Matches   88;  Conservative   46;  Mismatches  172;  Indels  104;  Gaps   16;

Qy       2 NIKKTAVKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSS 61
           |||   |:  :||                |  :|  | |               | ||
Db     171 NIKSLLVEVNIAVE--------------KGGSVYAKERVSLEN--------VTEATFSS 207

Qy      62 MFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAF 121
             ||  |   |  : | |: :   |   || |     |  |   |:   :      :
Db     208 -----NGGEQGGGGIYSEQDMLISDCNNVHFQGNAAGATAVKQCLDEEMIVLLAECVDSL 262

Qy     122 HTYGLEWTPNYVRWTVDGQE--VRKTEGGQVS--------------------------- 151
             |: ||   :   :| :    :||  |||
Db     263 SEDTLDSTPETEQTESNGNQDGSSETEDTQVSESPESTPSPDDVLGKGGGIYTEKSLTIT 322

Qy     152 NLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTD 211
           :|||     |: :::| |  | | :  |      | ::  | : ||  ||:  |   :
Db     323 GITGTIDFVSNI-ATDSGA--GVFTKENLSCTN-TNSLQFLKNSAGQHGGGAYVTQTMSV 378

Qy     212 NFDTFDGSRWGK--GDWTFDGNRVDLTDK--------NIYSRDGMLILALTRKGQESFNG 261
              | :        |:  |  |    | |         | |   :  : ||:   :   |
Db     379 TNTTSESITTPPLIGEVIFSEN----TAKGHGGGICTNKLSLSNLKTVTLTKNSAKESGG 434

Qy     262 -------QVPRDDEPAPQS-SSSAPASSSSVPAS--------SSSVPASSSSAFVPPSSS 305
                 :|  | |   : |||:|||:  | ||        |:: ||: |   :  : |
Db     435 AIFTDLASIPITDTPESSTPSSSSPASTPEVVASAKINRFFASTAKPAAPS---LTEAES 491

Qy     306 SATNAIHGMRTTPAVAKEHRNLVNA---------KGAKVNPNGHKRYRVN 346
             |:      |   :     |::|          ||  :     |  |:|
Db     492 DQTDQTETSDTNSDIDVSIENILNVAINQNTSAKKGGAIYGKKAKLSRIN 541


RESULT 9
US-09-801-368-362
; Sequence 362, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. US20020128250Alman, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 362
;   LENGTH: 526
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-801-368-362

  Query Match             5.3%;  Score 96.5;  DB 10;  Length 526;
  Best Local Similarity   21.0%;  Pred. No. 1.8;
  Matches   79;  Conservative   49;  Mismatches  147;  Indels  101;  Gaps   15;

Qy      36 YTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSN 95
           ||::     |   |:: | |   :  : :|| |        :::  :|    |  |
Db       6 YTMQGYNLVKLLKRLEEATA--RLEDVTIYQEG--------YIQNKLEASKNNKPS---- 51

Qy      96 IITGKAGAQKTSEKHHAVSPAADQ------AFHTYGLEWTPNYVRW-----TVDGQEVRK 144
                | |  |:|     :|  :|      || :|  |     |       ||    ::
Db      52 --DSGADANTTNEPSAENAPEVEQDPKCITAFQSYIGENIDPLVELSGKIDTVVLDALQL 109

Qy     145 TEGGQVSNLTGTQGL---RFNLWSSESAA----------WVGQFDESKLPLFQFINWVK 190
           :||  | ||  :     |   :||:: |           :||  ||           |
Db     110 LKGGFQSQLTFLRAAVRSRKPDYSSQTFADSLRPINENIIKLGQLKESN-------RQSK 162

Qy     191 VYKYTPGQGEGGSDFTLDWTDN-----FDTFDGSRWGKGDWT---------FDGNRVDLT 236
           : |      ||   |:    |       |  | :::    ||          | | |:
Db     163 YFAYLSALSEGAPLFSWVAVDTPVSMVTDFKDAAQF----WTNRILKEYRESDPNAVEWV 218

Qy     237 DKNIYSRDGMLIL--------ALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASS 288
           | : | | :              :|    |   : :  :    :|| :|||::: ||
Db     219 KKFLASFDNLKAYIKEYHTTGVSWKKDGMDFADAMAQSTKNTGATSSPSPASATAAPAPP 278

Qy     289 SSVPASSSSAF-----VPPSSSSAT--------------NAIHGMRTTPAVAKEHRN-- 326
             ||  :| |       | :|| |                |   |::      : |:|
Db     279 PPPPAPPASVFEISNDTPATSSDANKGGIGAVFAELNQGENITKGLKKVDKSQQTHKNPE 338

Qy     327 -----LVNAKGAKVNP 337
                 |:: |:|  |
Db     339 LRQSSTVSSTGSKSGP 354


RESULT 10
US-09-841-132-195
; Sequence 195, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 195
;   LENGTH: 821
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-841-132-195
```

```
RESULT 6
US-09-815-242-13423
; Sequence 13423, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13423
;   LENGTH: 719
;   TYPE: PRT
;   ORGANISM: Streptococcus pneumoniae
US-09-815-242-13423

  Query Match             5.8%;  Score 106.5;  DB 10;  Length 719;
  Best Local Similarity   23.5%;  Pred. No. 0.4;
  Matches   76;  Conservative   49;  Mismatches  105;  Indels   93;  Gaps   20;

Qy      57 GTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN------------PGSFQSNIITGKAGA 103
           | ::  :  |    :    | ||  :  :| |             |    || |:  :
Db     417 GNITLQYALQQSRNV----PAVET-LNKVGLNRAKTFLNGLGIDYPSIHYSNAIS--SNT 469

Qy     104 QKTSEKHHAVS---PAADQAFHTYGLEWTPNYVRWTV--DGQEVRKTEGGQVSNLTGTQG 158
           ::.:|: | |    ||  ||   |  : | |:   |  || |       : ||: ||:
Db     470 TESDKKYGASSEKMAAAYAAFANGGTYYKPMYIHKVVFSDGSE------KEFSNV-GTRA 522

Qy     159 LRFNLWSSESAAWVGQFDESKLPLF------QFINWVKVYKYTPGQGE-GGSDFTLDWTD 211
           ::      |: |::   |  | |        :: |:      |  |: | |::| :  :
Db     523 MK------ETTAYM-MTDMMKTVLSYGTGRNAYLAWL------PQAGKTGTSNYTDEEIE 569

Qy     212 NF----------DTFDG--SRWGKGDWTFDGNRV------DLTDKNIYSRDGMLILA--- 250
           |          : | |   ::    ||   ||:       ||      |  |  |:
Db     570 NHIKTSQFVAPDELFAGYTRKYSMAVWTGYSNRLTPLVGNGLTVAAKVYRSMMTYLSEGS 629

Qy     251 ----------LTRKGQESF-NGQVPRDDEPAPQ----SSSSAPASSSSVPASSSSVPASS 295
                    | | |:  | ||       ||||    : ||: :| ||   |||: |:::
Db     630 NPEDWNIPEGLYRNGEFVFKNGARSTWSSPAPQQPPSTESSSSSSSDSSTSQSSSTTPSTN 689

Qy     296 SSAFVPPSSSSATNAIHGMRTTP 318
           :|    |::::  :      |||
Db     690 NSTTTNPNNNTQQS-----NTTP 707


RESULT 7
US-09-738-363-2
; Sequence 2, Application US/09738363
; Patent No. US20010010932A1
;  GENERAL INFORMATION:
;       APPLICANT: Schnepf, Harry E.
;                  Schwab, George E.
;                  Payne, Jewel M.
;                  Narva, Kenneth E.
;                  Foncerrada, Luis
;       TITLE OF INVENTION: Nematicidal Proteins
;       NUMBER OF SEQUENCES: 40
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Jay M. Sanders
;            STREET: 2421 N.W. 41st Street, Suite A-1
;            CITY: Gainesville
;            STATE: FL
;            COUNTRY: USA
;            ZIP: 32606
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/738,363
;            FILING DATE: 15-Dec-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/076,137
;            FILING DATE: 12-MAY-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Sanders, Jay
;            REGISTRATION NUMBER: 39,355
;            REFERENCE/DOCKET NUMBER: MA-20CCCD3
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 352-375-8100
;            TELEFAX: 352-372-5800
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1385 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: YES
;       ANTI-SENSE: NO
;       ORIGINAL SOURCE:
;            ORGANISM: BACILLUS THURINGIENSIS
;            INDIVIDUAL ISOLATE: PS17
;       IMMEDIATE SOURCE:
;            CLONE: E. coli NM522(pMYC 1627) NRRL B-18651
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-738-363-2

  Query Match             5.3%;  Score 97.5;  DB 10;  Length 1385;
  Best Local Similarity   23.7%;  Pred. No. 5.5;
  Matches   59;  Conservative   36;  Mismatches   91;  Indels   63;  Gaps   13;

Qy      56 SGTVSSMFLYQNGSEIA-----DGRPWVEVD-IEVLGKNPGSFQSNIITGKAGAQKTSEK 109
           :|:: |   |||   |        | |   |      |     ||  :   |  ::::
Db    1176 NGSIRSDISYQNIDAIVLPTLPKLRHWFMSDRFSEQGDIMAKFQGALNRAYAQLEQSTLL 1235

Qy     110 HHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESA 169
           |:               :| :   ||::|      :  |::   | : ||   |||
Db    1236 HNG-------------HFTKDAANWTIEG------DAHQITLEDGRRVLRLPDWSSS-- 1273

Qy     170 AWVGQFDESKLPLFQFINWVKVYKYT-PGQGEGGSDFTLD----------WTDNFDTFDG 218
             | |  |      :  |  | |     |||||    ||:           | :|  |
Db    1274 --VSQMIE-----IENFNPDKEYNLVFHGQGEG--TVTLEHGEETKYIETHTHHFANFTT 1324

Qy     219 SRWGKGDWTFDGNRVDLTDKNIYSRDGMLI---LALTRKGQESFNGQVPRDDEPAPQSSS 275
           |:  :   ||: |:| :|   | | ||  :   :||          :| ||: :  :::
Db    1325 SQ--RQGLTFESNKVTVT---ISSEDGEFLVDNIALV-------EAPLPTDDQNSEGNTA 1372

Qy     276 SAPASSSSV 284
```

```
          :|| :    :|:|   : |:|:  |       |     :  ::    :  : |
Db    142 WLYSDFDYSVANGETVYSEIDVVELQQFDWYEGHQDDIYDMDLNLHAVVKENGQGVWKRP 201

Qy    103 ---AQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEG---GQVSNLTGT 156
            |:   |  |: |: |  || || |   | : | ||| || :       :  |:| :
Db    202 KMYPQEQLNKWRAMDPSKD--FHIYGCEVNQNEIIWYVDGVEVARKPNKYWHRPMNVTLS 259

Qy    157 QGLR---FNLWSSESAAWVGQFDE------SKLPLFQFINWVKVYKYTPGQGEGGSDFTL 207
           |||      : ::: |   : |       | :|   ::::|:|:: : |
Db    260 LGLRKPFVKFFDNKNNAINPETDAKAREKLSDIPTSMYVDYVRVWEKSAGN--------- 310

Qy    208 DWTDNFDTFD-------GSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFN 260
            | |  | :        ||:   |     :     |           |: ||  :
Db    311 --TTNPPTSEVGTLKTKGSKLVIDHWDASTGTISAVSNN-----------TKTGQ--YA 354

Qy    261 GQVPRDDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFVPPSS-----SSATN 309
          | |          :::: |   :: |::|   ::   |  | :|     | |:|
Db    355 GSV----------NNASIAQIVTLKANTSYKVSAFGKASSPGTSAYLGISKASN 398


RESULT 4
US-09-841-132-353
; Sequence 353, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 353
;   LENGTH: 583
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-841-132-353

  Query Match             5.9%;  Score 108;  DB 10;  Length 583;
  Best Local Similarity   22.4%;  Pred. No. 0.23;
  Matches   89;  Conservative   51;  Mismatches  174;  Indels   84;  Gaps   19;

Qy      8 VKSALAVAAAAAALTTNVSAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQN 67
          :  |:|:|      | ::    | |     | |      ||::    |   :|: :
Db     27 IGQAMAIAGQIKLPTVHIGPTAFLG-----LGVVDNNGNGARVQRVVGSAPAASLGI-ST 80

Qy     68 GSEI--ADGRP---WVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFH 122
          |  |   || |      :   : | :||   |     |:|  :|     |  | |:  |
Db     81 GDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNVTLAEGPPAE--FC 138

Qy    123 TYGLEWTPNYVRWTVDGQEVRKTE-----------GG-------QVSNLTGTQGLRFNL 163
          |   | |      :| |     |              ||        :: :|||     |:
Db    139 RYPSHWRP------LDTQVSESPESTPSPDDVLGKGGGIYTEKSLTITGITGTIDFVSNI 192

Qy    164 WSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNFDTFDGSRWGK 223
          :::| |  | | :  |      | ::  | : ||  ||:  |   :    | :
Db    193 -ATDSGA--GVFTKENLSCTN-TNSLQFLKNSAGQHGGGAYVTQTMSVTNTTSESITTPP 248

Qy    224 --GDWTFDGNRVDLTDK--------NIYSRDGMLILALTRKGQESFNG-------QVPRD 266
            |:  |  |    | |        |  |   :  : ||:   :   |          :|
Db    249 LVGEVIFSEN----TAKGHGGGICTNKLSLSNLKTVTLTKNSAKESGGAIFTDLASIPTT 304

Qy    267 DEPAPQS-SSSAPASSSSVPAS--------SSSVPASSSSAFVPPSSSATNAIHGMRTT 317
          | |   : |||:|||:  | ||        |:: ||: |   :  : |  |:      |
Db    305 DTPESSTPSSSSPASTPEVVASAKINRFFASTAEPAAPS---LTEAESDQTDQTETSDTN 361

Qy    318 PAVAKEHRNLVNA---------KGAKVNPNGHKRYRVN 346
            :    |::|          ||  :     |  |:|
Db    362 SDIDVSIENILNVAINQNTSAKKGGAIYGKKAKLSRIN 399
```

```
RESULT 5
US-09-765-272-2
; Sequence 2, Application US/09765272
; Patent No. US20020061545A1
;  GENERAL INFORMATION:
;       APPLICANT: Choi et. al.
;       TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;       NUMBER OF SEQUENCES: 452
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/765,272
;            FILING DATE: 22-Jan-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/961,083
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Brookes, A. Anders
;            REGISTRATION NUMBER: 36,373
;            REFERENCE/DOCKET NUMBER: PB340P2
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (301) 309-8504
;            TELEFAX: (301) 309-8512
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 666 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-272-2

  Query Match             5.9%;  Score 107.5;  DB 10;  Length 666;
  Best Local Similarity   23.5%;  Pred. No. 0.3;
  Matches   76;  Conservative   50;  Mismatches  104;  Indels   93;  Gaps   20;

Qy     57 GTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN------------PGSFQSNIITGKAGA 103
          | ::   :  |    :    | ||  :  :| |            |   || |:  :
Db    364 GNITLQYALQQSRNV----PAVET-LNKVGLNRAKTFLNGLGIDYPSIHYSNAIS--SNT 416

Qy    104 QKTSEKHHAVS---PAADQAFHTYGLEWTPNYVRWTV--DGQEVRKTEGGQVSNLTGTQG 158
           :: :|: | |    ||  ||   |  : | |:   |  || |       : ||: ||:
Db    417 TESDKKYGASSEKMAAAYAAFANGGTYYKPMYIHKVVFSDGSE------KEFSNV-GTRA 469

Qy    159 LRFNLWSSESAAWVGQFDESKLPLF------QFINWVKVYKYTPGQGE-GGSDFTLDWTD 211
          ::      |: |::   |  |  |         :: |:       |  |: | |::|':  :
Db    470 MK------ETTAYM-MTDMMKTVLTYGTGRNAYLAWL------PQAGKTGTSNYTDEEIE 516

Qy    212 NF----------DTFDG--SRWGKGDWTFDGNRV------DLTDKNIYSRDGMLILA--- 250
          |           : | |   ::    ||   ||:        ||      |  |  |:
Db    517 NHIKTSQFVAPDELFAGYTRKYSMAVWTGYSNRLTPLVGNGLTVAAKVYRSMMTYLSEGS 576

Qy    251 ----------LTRKGQESF-NGQVPRDDEPAPQ----SSSSAPASSSSVPASSSSVPASS 295
                    | | |:  | ||     : ||||    : ||: :| ||  |||: |:::
Db    577 NPEDWNIPEGLYRNGEFVFKNGARSTWNSPAPQQPPSTESSSSSSDSSTSQSSSTTPSTN 636

Qy    296 SSAFVPPSSSSATNAIHGMRTTP 318
          :|    |:::  :      |||
Db    637 NSTTTNPNNNTQQS-----NTTP 654
```

```
Qy    157 ----------QGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFT 206
                  |:  ::|   | |  |     ::|    |||               : ||
Db    175 PGTKYLNQQAMGVYISIWDGSSWATQG----GRVP----INW-----------ASAPFT 214

Qy    207 LDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKN 239
           :              |:  :|  ||  | |
Db    215 ATYQ--------------DFALNGCVVDPNDPN 233


RESULT 2
US-09-988-200-6
; Sequence 6, Application US/09988200
; Patent No. US20020094553A1
;   GENERAL INFORMATION:
;        APPLICANT: BARBEYRON, Tristan
;                   POTIN, Philippe
;                   RICHARD, Christophe
;                   HENRISSAT, Bernard
;                   YVIN, Jean-Claude
;                   KLOAREG, Bernard
;        TITLE OF INVENTION: Glycolyse hydrolase genes and their
;                            use for producing enzymes for the biodegradation of
;                            carrageenans
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DENNISON, MESEROLE, SCHEINER & SCHULTZ
;             STREET: 612 Crystal Square 4, 1745 Jefferson Davis
;                     Highway
;             CITY: ARLINGTON
;             STATE: VIRGINIA
;             COUNTRY: U.S.A.
;             ZIP: 22202
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/988,200
;             FILING DATE: 19-No. US20020094553A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/269,731
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: FR 96 12204
;             FILING DATE: 07-OCT-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: IRA SCHULTZ
;             REGISTRATION NUMBER: <Unknown>
;             REFERENCE/DOCKET NUMBER: <Unknown>
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 412-1155
;             TELEFAX: (703) 412-1161
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 620 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-988-200-6

  Query Match             6.4%;  Score 117;  DB 10;  Length 620;
  Best Local Similarity   27.3%;  Pred. No. 0.044;
  Matches   54;  Conservative   23;  Mismatches   73;  Indels   48;  Gaps   10;

Qy     43 YGKFEARMKMAAASGTVSSMF-LYQ--NGSEIADGR-PWVEVDI---------------- 82
          || :|||:| |:    ||  | :|   : |   :|   : |:|:
Db    289 YGYYEARIKGASTFPGVSPAFWMYSTIDRSLTKEGDVQYSEIDVVELTQKSAVRESDHDL 348

Qy     83 -EVLGKN-------PGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVR 134
           :: ||       ||||      |          |    |  |  |||||:  | : :
```

```
Db    349 HNIVVKNGKPTWMRPGSFPQTNHNG---------YHLPFDPRND--FHTYGVNVTKDKIT 397

Qy    135 WTVDGQEVRKTEG---GQVSNLTGTQGLR--FNLWSSE----SAAWVGQFDESKLPLFQF 185
          | |||: | : :     :  ||| :||||     |       ||    :   : : :
Db    398 WYVDGEIVGEKDNLYWHRQMNLTLSQGLRAPHTQWKCNQFYPSANKSAEGFPTSMEVDYV 457

Qy    186 INWVKVYKYTPGQGEGGS 203
            ||||       ||| |
Db    458 RTWVKVGNNNSAPGEGQS 475


RESULT 3
US-09-988-200-8
; Sequence 8, Application US/09988200
; Patent No. US20020094553A1
;   GENERAL INFORMATION:
;        APPLICANT: BARBEYRON, Tristan
;                   POTIN, Philippe
;                   RICHARD, Christophe
;                   HENRISSAT, Bernard
;                   YVIN, Jean-Claude
;                   KLOAREG, Bernard
;        TITLE OF INVENTION: Glycolyse hydrolase genes and their
;                            use for producing enzymes for the biodegradation of
;                            carrageenans
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DENNISON, MESEROLE, SCHEINER & SCHULTZ
;             STREET: 612 Crystal Square 4, 1745 Jefferson Davis
;                     Highway
;             CITY: ARLINGTON
;             STATE: VIRGINIA
;             COUNTRY: U.S.A.
;             ZIP: 22202
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/988,200
;             FILING DATE: 19-No. US20020094553A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/269,731
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: FR 96 12204
;             FILING DATE: 07-OCT-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: IRA SCHULTZ
;             REGISTRATION NUMBER: <Unknown>
;             REFERENCE/DOCKET NUMBER: <Unknown>
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 412-1155
;             TELEFAX: (703) 412-1161
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 545 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-988-200-8

  Query Match             6.3%;  Score 114.5;  DB 10;  Length 545;
  Best Local Similarity   20.6%;  Pred. No. 0.06;
  Matches   73;  Conservative   61;  Mismatches  131;  Indels   89;  Gaps   17;

Qy      8 VKSALAVAAAAAALTTNVSAKD----FSGAELYTLEEVQYGKFEARMKMA-AASGTVSSM 62
          || :  :|        | :  |    |:     : ::  || |||::: |    |   |
Db     82 VKISNGIAELTMRHNANNTPPDGGTYFTSGIFKSYQKFTYGYFEAKIQGADIGEGVCPSF 141

Qy     63 FLYQN-GSEIADGRP-WVEVDIEVL-------GKNPGSFQSNI----ITGKAG------- 102
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:16 ; Search time 8.88657 Seconds
(without alignments)
761.932 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKKTAVKSALAVAAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	120	6.6	287	10	US-09-734-569-152	Sequence 152, App
2	117	6.4	620	10	US-09-988-200-6	Sequence 6, Appli
3	114.5	6.3	545	10	US-09-988-200-8	Sequence 8, Appli
4	108	5.9	583	10	US-09-841-132-353	Sequence 353, App
5	107.5	5.9	666	10	US-09-765-272-2	Sequence 2, Appli
6	106.5	5.8	719	10	US-09-815-242-13423	Sequence 13423, A
7	97.5	5.3	1385	10	US-09-738-363-2	Sequence 2, Appli
8	97	5.3	1770	10	US-09-841-132-444	Sequence 444, App
9	96.5	5.3	526	10	US-09-801-368-362	Sequence 362, App
10	96	5.3	821	10	US-09-841-132-195	Sequence 195, App
11	96	5.3	1776	10	US-09-841-132-179	Sequence 179, App
12	95.5	5.2	1289	10	US-09-738-363-4	Sequence 4, Appli
13	93	5.1	156	10	US-09-925-301-1154	Sequence 1154, Ap
14	93	5.1	714	9	US-09-738-626-6377	Sequence 6377, Ap
15	91	5.0	279	9	US-09-738-626-3942	Sequence 3942, Ap
16	91	5.0	678	10	US-09-801-368-314	Sequence 314, App
17	91	5.0	1751	10	US-09-841-132-445	Sequence 445, App
18	91	5.0	1751	10	US-09-841-132-594	Sequence 594, App
19	90	4.9	674	9	US-10-086-464-14	Sequence 14, Appl
20	89.5	4.9	493	9	US-09-738-626-6780	Sequence 6780, Ap
21	89.5	4.9	1162	10	US-09-745-008-34	Sequence 34, Appl
22	89	4.9	499	12	US-10-052-586-592	Sequence 592, App
23	89	4.9	1367	10	US-09-801-368-108	Sequence 108, App
24	88.5	4.9	630	10	US-09-815-242-10036	Sequence 10036, A
25	88	4.8	1723	10	US-09-841-132-394	Sequence 394, App
26	88	4.8	1723	10	US-09-841-132-395	Sequence 395, App
27	87.5	4.8	467	10	US-09-863-547B-1	Sequence 1, Appli
28	87.5	4.8	2167	10	US-09-778-927A-61	Sequence 61, Appl
29	87.5	4.8	2201	12	US-10-100-912-2	Sequence 2, Appli
30	87	4.8	655	10	US-09-205-658-57	Sequence 57, Appl
31	87	4.8	655	10	US-09-844-353A-57	Sequence 57, Appl
32	87	4.8	1884	10	US-09-785-770A-17	Sequence 17, Appl
33	87	4.8	1907	10	US-09-785-770A-16	Sequence 16, Appl
34	86.5	4.7	765	10	US-09-975-326-4	Sequence 4, Appli
35	86.5	4.7	778	10	US-09-815-242-10728	Sequence 10728, A
36	86.5	4.7	1463	9	US-09-971-536-69	Sequence 69, Appl
37	86	4.7	498	10	US-09-925-302-475	Sequence 475, App
38	86	4.7	671	9	US-09-792-630-43	Sequence 43, Appl
39	86	4.7	671	9	US-10-080-376-43	Sequence 43, Appl
40	86	4.7	766	9	US-09-934-406-2	Sequence 2, Appli
41	86	4.7	766	10	US-09-975-326-2	Sequence 2, Appli
42	86	4.7	871	10	US-09-886-468-21	Sequence 21, Appl
43	85.5	4.7	341	12	US-10-072-152-4	Sequence 4, Appli
44	85.5	4.7	848	10	US-09-841-132-192	Sequence 192, App
45	85.5	4.7	1530	10	US-09-841-132-178	Sequence 178, App

ALIGNMENTS

RESULT 1
US-09-734-569-152
; Sequence 152, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv
; TITLE OF INVENTION: in the synthesis of carbohydrates
; FILE REFERENCE: BASF-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734,569
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 152
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-152

Query Match 6.6%; Score 120; DB 10; Length 287;
Best Local Similarity 23.0%; Pred. No. 0.0089;
Matches 49; Conservative 30; Mismatches 70; Indels 64; Gaps 11;

```
Qy      43 YGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGS---FQSNII 97
           |    | :||    ::|||:: ::   | :       |:|:| ||   |      :|:
Db      69 YVDISAYIKMPPFDSAGTVTTFYMSSQGDQ------HYELDMEFLGNTSGQPFLLHTNVF 122

Qy      98 TGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGT- 156
             | ::  : :   |:|| || |  |: : | : || : ||          || ||
Db     123 VDGVGGRE-QQMYLGFDPSAD--FHYYRFRWSKDMVVFYVDNKPVR-----VFKNLEGTV 174
```

```
;  PRIOR APPLICATION NUMBER: 60/329,709
;  PRIOR FILING DATE: 2001-10-16
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 261
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
PCT-US02-33165-3

  Query Match             5.4%;  Score 98;  DB 1;  Length 261;
  Best Local Similarity   21.9%;  Pred. No. 20;
  Matches   70;  Conservative   36;  Mismatches  100;  Indels  114;  Gaps   17;

Qy      52 MAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHH 111
           :||         | |:|:|:               | |||           : : | |
Db      15 LAAQPAMAEVQLLEQSGAEV--------------KKPGS----------SVQVSCK-- 46

Qy     112 AVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAW 171
             |:   |  ||  |              ||:  |          ||          |
Db      47 ----ASGGTFSMYGFNW------------VRQAPG---------HGLE----------W 70

Qy     172 VGQFDESKLPLFQFINWVKVYK----YTPGQGEG-----------------GSDFTLD 208
           :|      :|:|   |: : ::    :|  |                    ||  |
Db      71 MGGI----IPIFGTSNYAQKFRGRVTFTADQATSTAYMELTNLRSDDTAVYYCARDFGPD 126

Qy     209 WTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQ-----V 263
           | |  |::|||  |:| : | |    :|  :  |  |  :  |    : :  |      :
Db     127 WEDG-DSYDGS--GRGFFDFWGQGTLVTVSSA-STKGPSVFPLAPSSKSTSGGTAALGCL 182

Qy     264 PRDDEPAPQSSS-SAPASSSSV---PA--SSSSVPASSSSAFVPPSSSSATNAIHGMRTT 317
            :|  | | : | :: | :| |   ||   || : : ||   || ||      |  :
Db     183 VKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHK 242

Qy     318 PAVAKEHRNLVNAKGAKVNP 337
           |:  |  :        || |
Db     243 PSNTKVDK--------KVEP 254


Search completed: January  9, 2003, 12:14:20
Job time : 22.3495 secs
```

```
RESULT 12
US-09-724-676A-83344
; Sequence 83344, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83344
;   LENGTH: 1602
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83344

  Query Match             5.4%;  Score 98.5;  DB 5;  Length 1602;
  Best Local Similarity   19.5%;  Pred. No. 1.4e+02;
  Matches   51;  Conservative   33;  Mismatches   85;  Indels   93;  Gaps    9;

Qy      35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
           | |:   :|  :|: ::    |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
           : ::   |   ::  |  :|       || |   ||: | |     : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     154 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 173
            |         :||       |    | |   |||
Db    1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy     174 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 200
                                    ::| |  |: : :| :     :|||        |
Db    1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy     201 GGSDFTLDWTDNFDTFDGSRWG 222
           | | |:| :    |||:|   |
Db    1431 GKSSFSLAFFRMVDTFEGQLEG 1452


RESULT 13
US-09-724-676-83342
; Sequence 83342, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83342
;   LENGTH: 1629
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83342

  Query Match             5.4%;  Score 98.5;  DB 5;  Length 1629;
  Best Local Similarity   19.0%;  Pred. No. 1.4e+02;
  Matches   60;  Conservative   45;  Mismatches  108;  Indels  103;  Gaps   12;

Qy      35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
           | |:   :|  :|: ::    |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
           : ::   |   ::  |  :|       || |   ||: | |     : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     154 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 173
            |         :||       |    | |   |||
Db    1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy     174 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 200
                                    ::| |  |: : :| :     :|||        |
Db    1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy     201 GGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFN 260
           | | |:| :    |||:       |    ||  :|:    :::    | : |  :    |:
Db    1431 GKSSFSLAFFRMVDTFE------GHIIIDG--IDIAKLPLHTLRSRLSIIL--QDPVLFS 1480

Qy     261 GQVPRDDEPAPQSSSS 276
           | :  : :|  : | |
Db    1481 GTIRFNLDPERKCSDS 1496


RESULT 14
US-09-724-676A-83342
; Sequence 83342, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83342
;   LENGTH: 1629
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83342

  Query Match             5.4%;  Score 98.5;  DB 5;  Length 1629;
  Best Local Similarity   19.0%;  Pred. No. 1.4e+02;
  Matches   60;  Conservative   45;  Mismatches  108;  Indels  103;  Gaps   12;

Qy      35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
           | |:   :|  :|: ::    |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
           : ::   |   ::  |  :|       || |   ||: | |     : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     154 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 173
            |         :||       |    | |   |||
Db    1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy     174 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 200
                                    ::| |  |: : :| :     :|||        |
Db    1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy     201 GGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFN 260
           | | |:| :    |||:       |    ||  :|:    :::    | : |  :    |:
Db    1431 GKSSFSLAFFRMVDTFE------GHIIIDG--IDIAKLPLHTLRSRLSIIL--QDPVLFS 1480

Qy     261 GQVPRDDEPAPQSSSS 276
           | :  : :|  : | |
Db    1481 GTIRFNLDPERKCSDS 1496


RESULT 15
PCT-US02-33165-3
; Sequence 3, Application PC/TUS0233165
; GENERAL INFORMATION:
;  APPLICANT: GOVERNMENT OF THE UNITED STATES OF AMERICA, REPRESENTED BY THE SECRETAR
;  APPLICANT:  DEPARTMENT OF HEALTH AND HUMAN SERVICES
;  APPLICANT:  THE SCRIPPS RESEARCH INSTITUTE
;  TITLE OF INVENTION: BROADLY CROSS-REACTIVE NEUTRALIZING ANTIBODIES AGAINST HUMAN
;  TITLE OF INVENTION:  IMMUNODEFICIENCY VIRUS SELECTED BY ENV-CD4-CO-RECEPTOR COMPLE
;  FILE REFERENCE: 218837
;  CURRENT APPLICATION NUMBER: PCT/US02/33165
;  CURRENT FILING DATE:  2002-10-16
```

```
          : | | |  |  |     :|    :    |: :|  ||   | |         : : :||
Db    102 VDTYKIGIDKAIEAAQKSTPTL-LSLMDEGVRFTNGYVAHGVSGP-------SRAAIMTG 153

Qy    156 TQGLRFNLWSSESAAWVGQFDESKLP-LFQFINWVKVYKYTPGQGE----------GGS 203
              || ::|:  |       |: || |||   :     ||   |:
Db    154 RAPARFGVYSNTDAQDGIPLTETFLPELFQNHGY-----YTAAVGKWHLSKISNVPVPED 208

Qy    204 DFTLDWTDNFDTFDGSRWGKGDWTFD 229
            | |: ||| ||    |   :  ||
Db    209 KQTRDYHDNFTTFSAEEWQPQNRGFD 234


RESULT 9
US-09-724-676-83343
; Sequence 83343, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83343
;   LENGTH: 1558
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83343

  Query Match             5.4%;  Score 99;  DB 5;  Length 1558;
  Best Local Similarity   18.7%;  Pred. No. 1.2e+02;
  Matches   59;  Conservative   38;  Mismatches  100;  Indels  118;  Gaps   11;

Qy     35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    154 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 173
            |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    174 -------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 200
                                      ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    201 GGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR-----------------------VDL 235
          | | |:| :    |||:|    :|   |  | |                        :|:
Db   1431 GKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARAFVRKTSIFIMDEATASIDM 1490

Qy    236 TDKNIYSRDGMLILA 250
            :||  :  |   |
Db   1491 ATENILQKVVMTAFA 1505


RESULT 10
US-09-724-676A-83343
; Sequence 83343, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83343
;   LENGTH: 1558
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83343

  Query Match             5.4%;  Score 99;  DB 5;  Length 1558;
  Best Local Similarity   18.7%;  Pred. No. 1.2e+02;
  Matches   59;  Conservative   38;  Mismatches  100;  Indels  118;  Gaps   11;

Qy     35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    154 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 173
            |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    174 -------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 200
                                      ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    201 GGSDFTLDWTDNFDTFDGSRWG 222
          | | |:| :    |||:|    |
Db   1431 GKSSFSLAFFRMVDTFEGQLEG 1452


US-09-724-676A-83343

  Query Match             5.4%;  Score 99;  DB 5;  Length 1558;
  Best Local Similarity   18.7%;  Pred. No. 1.2e+02;
  Matches   59;  Conservative   38;  Mismatches  100;  Indels  118;  Gaps   11;

Qy     35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    154 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 173
            |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    174 -------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 200
                                      ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    201 GGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR-----------------------VDL 235
          | | |:| :    |||:|    :|   |  | |                        :|:
Db   1431 GKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARAFVRKTSIFIMDEATASIDM 1490

Qy    236 TDKNIYSRDGMLILA 250
            :||  :  |   |
Db   1491 ATENILQKVVMTAFA 1505


RESULT 11
US-09-724-676-83344
; Sequence 83344, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83344
;   LENGTH: 1602
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83344

  Query Match             5.4%;  Score 98.5;  DB 5;  Length 1602;
  Best Local Similarity   19.5%;  Pred. No. 1.4e+02;
  Matches   51;  Conservative   33;  Mismatches   85;  Indels   93;  Gaps    9;

Qy     35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    154 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 173
            |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    174 -------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 200
                                      ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    201 GGSDFTLDWTDNFDTFDGSRWG 222
          | | |:| :    |||:|    |
Db   1431 GKSSFSLAFFRMVDTFEGQLEG 1452
```

```
;   LENGTH: 771
;   TYPE: PRT
;   ORGANISM: Lactobacillus rhamnosus
US-10-264-213-165

  Query Match             5.7%;  Score 103.5;  DB 6;  Length 771;
  Best Local Similarity   22.7%;  Pred. No. 28;
  Matches   79;  Conservative   37;  Mismatches  117;  Indels  115;  Gaps   18;

Qy      32 GAELYTLEEV-QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGR---------------- 75
           || : |:||   || |       |  ||    : | |    |||
Db     493 GASVNTVEEAGAYGAF-------ANGGTYYKPY-YVNKVVSADGNTQTFNSQGTRAMKSS 544

Qy      76 -PWVEVDI--EVLGKNPGS-------FQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYG 125
             ::  |:   || |  |:       :|:    || |    |::    :|    | :  |
Db     545 TAYMITDMLKGVLTKGTGTSAAISGLYQA----GKTGTTDYSDEELKQNP----ALNATG 596

Qy     126 L---EWTPNYVR------WTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFD 176
           :    |   | |       ||  | :   : |   :  | :|
Db     597 IAKDAWFTGYTRNRVISVWT--GYDKPTSHGISYAEQTISQ------------------- 635

Qy     177 ESKLPLFQFINWVKVYK----YTPGQGEGGSDFTL-DWTDNFDTFDGSRWGKGDWTFDGN 231
                     |:||    ||  |     |:|  |  ::::   ||  |    :   |
Db     636 ------------KIYKALMSYT-SQNLDNKDWTKPDTVESYNILKGSNPGTAITSGSAN 681

Qy     232 RVDLTDKNIYSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSV 291
               | | :| |                 |  |   :   :||||: ||| |   |||
Db     682 ----TTKELYVR----------------GHGPSSHKAVAESSSSSSASSES---SSSKE 717

Qy     292 PASSSSAFVPPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNG 339
            :| ||  |  | :|   :     | |: |   :   ::    |:  |
Db     718 SSSESSNAVSASIASHATSSSVSSTAPSSAAPSKEETSSSSEPVSSAG 765


RESULT 6
US-09-724-676-83341
; Sequence 83341, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83341
;   LENGTH: 1554
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83341

  Query Match             5.7%;  Score 103.5;  DB 5;  Length 1554;
  Best Local Similarity   20.8%;  Pred. No. 61;
  Matches   46;  Conservative   36;  Mismatches   80;  Indels   59;  Gaps    8;

Qy      35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
           : ::   |    ::  |  :|       || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     154 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 189
            |       |  :|: :: |                      ::| |  |: : :| :
Db    1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy     190 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWG 222
               :|||        | | | |:| :    |||:|   |
Db    1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGQLEG 1404


RESULT 7
US-09-724-676A-83341
; Sequence 83341, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83341
;   LENGTH: 1554
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83341

  Query Match             5.7%;  Score 103.5;  DB 5;  Length 1554;
  Best Local Similarity   20.8%;  Pred. No. 61;
  Matches   46;  Conservative   36;  Mismatches   80;  Indels   59;  Gaps    8;

Qy      35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
           : ::   |    ::  |  :|       || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     154 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 189
            |       |  :|: :: |                      ::| |  |: : :| :
Db    1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy     190 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWG 222
               :|||        | | | |:| :    |||:|   |
Db    1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGQLEG 1404


RESULT 8
US-10-287-274-406
; Sequence 406, Application US/10287274
; GENERAL INFORMATION:
;  APPLICANT: Forsyth, R. Allyn
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THER
;  FILE REFERENCE: ELITRA.008DV1
;  CURRENT APPLICATION NUMBER: US/10/287,274
;  CURRENT FILING DATE:  2002-10-31
;  PRIOR APPLICATION NUMBER: US 60/164415
;  PRIOR FILING DATE: 1999-11-09
;  PRIOR APPLICATION NUMBER: US 09/711164
;  PRIOR FILING DATE: 2000-11-09
;  NUMBER OF SEQ ID NOS: 469
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 406
;   LENGTH: 571
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-287-274-406

  Query Match             5.4%;  Score 99;  DB 6;  Length 571;
  Best Local Similarity   21.8%;  Pred. No. 40;
  Matches   58;  Conservative   34;  Mismatches   88;  Indels   86;  Gaps   11;

Qy       3 IKKTAVKSAL---------AVAAAAA------ALTTNVSAKDFSGAE----------LYT 37
           :||: | :::          | || ||      |  |||:  ||:  |           : |
Db      16 LKKSVVSTSISLILASGMAAFAAHAADDVKLKATKTNVAFSDFTPTEYSTKGKPNIIVLT 75

Qy      38 LEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSN-- 95
           :::: ||:                    :  ||                   :| : ::
Db      76 MDDLGYGQLP----------------FDKGS----------------FDPKTMENREV 101

Qy      96 IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTG 155
```

```
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83340
;   LENGTH: 1510
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83340

  Query Match             5.7%;  Score 104;  DB 5;  Length 1510;
  Best Local Similarity   19.7%;  Pred. No. 55;
  Matches   54;  Conservative   41;  Mismatches   95;  Indels   84;  Gaps   10;

Qy      35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     154 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 189
           |       |  :|: :: |                        ::| |  |: : :| :
Db    1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy     190 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR-------- 232
             :|||        | | | |:| :    |||:|    :|   |  | |
Db    1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARA 1423

Qy     233 ----------------VDLTDKNIYSRDGMLILA 250
                           :|:  :||  :  |   |
Db    1424 FVRKTSIFIMDEATASIDMATENILQKVVMTAFA 1457


RESULT 3
US-09-724-676A-83340
; Sequence 83340, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83340
;   LENGTH: 1510
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83340

  Query Match             5.7%;  Score 104;  DB 5;  Length 1510;
  Best Local Similarity   19.7%;  Pred. No. 55;
  Matches   54;  Conservative   41;  Mismatches   95;  Indels   84;  Gaps   10;

Qy      35 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 93
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      94 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 153
           : ::   |    ::  |  :|        || |    ||: | |      : : | | ':
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     154 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 189
           |       |  :|: :: |                        ::| |  |: : :| :
Db    1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy     190 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR-------- 232
             :|||        | | | |:| :    |||:|    :|   |  | |
Db    1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARA 1423

Qy     233 ----------------VDLTDKNIYSRDGMLILA 250
                           :|:  :||  :  |   |
Db    1424 FVRKTSIFIMDEATASIDMATENILQKVVMTAFA 1457


RESULT 4
US-10-264-213-137
; Sequence 137, Application US/10264213
; GENERAL INFORMATION:
;  APPLICANT: Glenn, Matthew
;  APPLICANT:  Havukkala, Ilkka J
;  APPLICANT:  Lubbers, Mark William
;  APPLICANT:  Dekker, James
;  TITLE OF INVENTION: Polynucleotides, materials incorporating
;  TITLE OF INVENTION:  them, and methods for using them.
;  FILE REFERENCE: 11000.1043c3
;  CURRENT APPLICATION NUMBER: US/10/264,213
;  CURRENT FILING DATE:  2002-10-03
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 137
;   LENGTH: 771
;   TYPE: PRT
;   ORGANISM: Lactobacillus rhamnosus
US-10-264-213-137

  Query Match             5.7%;  Score 103.5;  DB 6;  Length 771;
  Best Local Similarity   22.7%;  Pred. No. 28;
  Matches   79;  Conservative   37;  Mismatches  117;  Indels  115;  Gaps   18;

Qy      32 GAELYTLEEV-QYGKFEARMKMAAASGTVSSMFLYQNGSEIADGR--------------- 75
           || : |:||   || |       |  ||    : | |    |||
Db     493 GASVNTVEEAGAYGAF-------ANGGTYYKPY-YVNKVVSADGNTQTFNSQGTRAMKSS 544

Qy      76 -PWVEVDI--EVLGKNPGS-------FQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYG 125
            ::  |:   || |  |:        :|:     || |    |::    :|    | :  |
Db     545 TAYMITDMLKGVLTKGTGTSAAISGLYQA----GKTGTTDYSDEELKQNP----ALNATG 596

Qy     126 L---EWTPNYVR------WTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFD 176
           :    |   | |      ||  | :   : |   :  | :|
Db     597 IAKDAWFTGYTRNRVISVWT--GYDKPTSHGISYAEQTISQ------------------- 635

Qy     177 ESKLPLFQFINWVKVYK----YTPGQGEGGSDFTL-DWTDNFDTFDGSRWGKGDWTFDGN 231
                      |:||    ||  |      |:|  |  ::::   ||  |    :   |
Db     636 -------------KIYKALMSYT-SQNLDNKDWTKPDTVESYNILKGSNPGTAITSGSAN 681

Qy     232 RVDLTDKNIYSRDGMLILALTRKGQESFNGQVPRDDEPAPQSSSSAPASSSSVPASSSSV 291
               | :| |                  |  |   :   :||||: ||| |   |||
Db     682 ----TTKELYVR----------------GHGPSSHKAVAESSSSSSASSES---SSSKE 717

Qy     292 PASSSSAFVPPSSSSATNAIHGMRTTPAVAKEHRNLVNAKGAKVNPNG 339
            :| ||  |  | :|   :     | |: |   :   ::    |:  |
Db     718 SSSESSNAVSASIASHATSSSVSSTAPSSAAPSKEETSSSSEPVSSAG 765


RESULT 5
US-10-264-213-165
; Sequence 165, Application US/10264213
; GENERAL INFORMATION:
;  APPLICANT: Glenn, Matthew
;  APPLICANT:  Havukkala, Ilkka J
;  APPLICANT:  Lubbers, Mark William
;  APPLICANT:  Dekker, James
;  TITLE OF INVENTION: Polynucleotides, materials incorporating
;  TITLE OF INVENTION:  them, and methods for using them.
;  FILE REFERENCE: 11000.1043c3
;  CURRENT APPLICATION NUMBER: US/10/264,213
;  CURRENT FILING DATE:  2002-10-03
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 165
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:04:12 ; Search time 15.3495 Seconds
(without alignments)
1613.076 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKKTAVKSALAVAAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 263044 seqs, 70945482 residues

Total number of hits satisfying chosen parameters: 263044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	116.5	6.4	673	6	US-10-294-561-3	Sequence 3, Appli
2	104	5.7	1510	5	US-09-724-676-83340	Sequence 83340, A
3	104	5.7	1510	5	US-09-724-676A-83340	Sequence 83340, A
4	103.5	5.7	771	6	US-10-264-213-137	Sequence 137, App
5	103.5	5.7	771	6	US-10-264-213-165	Sequence 165, App
6	103.5	5.7	1554	5	US-09-724-676-83341	Sequence 83341, A
7	103.5	5.7	1554	5	US-09-724-676A-83341	Sequence 83341, A
8	99	5.4	571	6	US-10-287-274-406	Sequence 406, App
9	99	5.4	1558	5	US-09-724-676-83343	Sequence 83343, A
10	99	5.4	1558	5	US-09-724-676A-83343	Sequence 83343, A
11	98.5	5.4	1602	5	US-09-724-676-83344	Sequence 83344, A
12	98.5	5.4	1602	5	US-09-724-676A-83344	Sequence 83344, A
13	98.5	5.4	1629	5	US-09-724-676-83342	Sequence 83342, A
14	98.5	5.4	1629	5	US-09-724-676A-83342	Sequence 83342, A
15	98	5.4	261	1	PCT-US02-33165-3	Sequence 3, Appli
16	98	5.4	1833	5	US-09-724-676-88017	Sequence 88017, A
17	98	5.4	1833	5	US-09-724-676A-88017	Sequence 88017, A
18	98	5.4	1841	5	US-09-724-676-88018	Sequence 88018, A
19	98	5.4	1841	5	US-09-724-676A-88018	Sequence 88018, A
20	98	5.4	1853	5	US-09-724-676-88014	Sequence 88014, A
21	98	5.4	1853	5	US-09-724-676A-88014	Sequence 88014, A
22	98	5.4	1860	5	US-09-724-676-88015	Sequence 88015, A
23	98	5.4	1860	5	US-09-724-676A-88015	Sequence 88015, A
24	98	5.4	1868	5	US-09-724-676-88016	Sequence 88016, A
25	98	5.4	1868	5	US-09-724-676A-88016	Sequence 88016, A
26	98	5.4	1880	5	US-09-724-676-88013	Sequence 88013, A
27	98	5.4	1880	5	US-09-724-676A-88013	Sequence 88013, A
28	96	5.3	1439	5	US-09-134-000C-6133	Sequence 6133, Ap
29	94	5.2	261	1	PCT-US02-33165-11	Sequence 11, Appl
30	91	5.0	270	6	US-10-276-774-1484	Sequence 1484, Ap
31	91	5.0	453	5	US-09-724-676-53371	Sequence 53371, A
32	91	5.0	453	5	US-09-724-676-53372	Sequence 53372, A
33	91	5.0	453	5	US-09-724-676A-53371	Sequence 53371, A
34	91	5.0	453	5	US-09-724-676A-53372	Sequence 53372, A
35	91	5.0	1121	5	US-09-917-384A-1	Sequence 1, Appli
36	90.5	5.0	328	6	US-10-171-404A-32	Sequence 32, Appl
37	90.5	5.0	1684	6	US-10-276-774-2398	Sequence 2398, Ap
38	89	4.9	467	5	US-09-724-676-56516	Sequence 56516, A
39	89	4.9	467	5	US-09-724-676A-56516	Sequence 56516, A
40	89	4.9	499	6	US-10-125-923A-592	Sequence 592, App
41	89	4.9	499	6	US-10-205-892-592	Sequence 592, App
42	89	4.9	499	6	US-10-174-575-592	Sequence 592, App
43	89	4.9	499	6	US-10-174-575A-592	Sequence 592, App
44	89	4.9	515	5	US-09-724-676-56515	Sequence 56515, A
45	89	4.9	515	5	US-09-724-676A-56515	Sequence 56515, A

ALIGNMENTS

RESULT 1
US-10-294-561-3
; Sequence 3, Application US/10294561
; GENERAL INFORMATION:
; APPLICANT: Seikagaku Corporation
; TITLE OF INVENTION: (1-3)-Beta-D-Glucan Binding Domain Protein, Measuring Method U
; TITLE OF INVENTION: Substance And Assay Kit
; FILE REFERENCE: Q72918
; CURRENT APPLICATION NUMBER: US/10/294,561
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: JP 2001-351943
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Tachypleus tridentatus
US-10-294-561-3

Query Match 6.4%; Score 116.5; DB 6; Length 673;
Best Local Similarity 22.3%; Pred. No. 3.3;
Matches 60; Conservative 41; Mismatches 97; Indels 71; Gaps 15;

```
Qy      21 LTTNVSAKDFSGAELYTLEEV--------QYGKFEARMKMAAASGTVSSMFLYQNGSEIA 72
           |       :|: | : ||    :         :||| ||:| : :  | |  ||
Db      72 LVITAKREDYDGFK-YTSARLKTQFDKSWKYGKIEAKMAIPSFRG-VWVMF--------- 120

Qy      73 DGRPWVEVDIEVLGKNPGSFQSNIITGK-AGAQKTSEKHHAVSPAADQA---------- 120
              |:  |     : | | : : |  :     :|     |  :|    |
Db     121 ----WMSGDNTNYVRWPSSGEIDFIEHRNTNNEKVRGTIHWSTPDGAHAHHNRESNTNGI 176

Qy     121 -FHTYGLEWTPNYVRWTVDGQEV--RKTEGGQVSNLTGTQGLRFNLWSSESAA----WVG 173
            :| | :||  : |:| |:| :    | :||    : |    |  ::   : |    | |
Db     177 DYHIYSVEWNSSIVKWFVNGNQYFEVKIQGG----VNGKSAFRNKVFVILNMAIGGNWPG 232

Qy     174 QFD--ESKLPLFQFINWVKVYKYTPGQGEGG-------------SDFTLDWTDNFDTFD 217
           ||  :   |   :|::|:||:         |              |:  || ||  :  |
Db     233 -FDVADEAFPAKMYIDYVRVYQDASTSSPVGDTSLDGYYFVQNRHSELYLDVTDASNE-D 290

Qy     218 GSRWGKGDWTFDGNR-----VDLTDKNIY 241
           |:   :  |:: ||       :  : |:|
Db     291 GAFLQQ--WSYSGNENQQFDFEHLENNVY 317
```

RESULT 2
US-09-724-676-83340
; Sequence 83340, Application US/09724676
; GENERAL INFORMATION:

Search completed: January 9, 2003, 12:13:26
Job time : 118.824 secs

```
US-09-791-537-31771
; Sequence 31771, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31771
;   LENGTH: 276
;   TYPE: PRT
;   ORGANISM: Bacillus sp
US-09-791-537-31771

  Query Match             18.6%;  Score 247.5;  DB 21;  Length 276;
  Best Local Similarity   36.1%;  Pred. No. 1.1e-16;
  Matches   61;  Conservative   20;  Mismatches   75;  Indels   13;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           :   || | :   || ||  || | ::||||| | |  |     :  || |:||| |||:
Db      96 YKAGELRTNDFYHYGLFEVSMKPAKSTGTVSSFFTY-TGPWDWENDPWDEIDIEFLGKDT 154

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |   |    :| :| :   |   |:||  || |  :|| |:|: |
Db     155 TKIQFNYFTNGVGG---NEHYHELGFDAADDFNTYAFEWRPESIRWFVNGELVHTA---- 207

Qy     127 VSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             |:  | | :  |||       | |:|:    |:    :||   ||||
Db     208 TENIPQTPQKIMMNLWPGIGVDGWTGRFNGEDTPVVTQYDWV---KYTP 253


RESULT 13
US-09-791-537-27063
; Sequence 27063, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27063
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-791-537-27063

  Query Match             18.5%;  Score 246.5;  DB 21;  Length 242;
  Best Local Similarity   35.5%;  Pred. No. 1.2e-16;
  Matches   60;  Conservative   24;  Mismatches   74;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :  || || |:||| |
Db      83 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFL 139

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
Db     140 GKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 195

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
             |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     196 ATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 14
US-09-791-537-73450
; Sequence 73450, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73450
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-791-537-73450

  Query Match             18.4%;  Score 245.5;  DB 21;  Length 242;
  Best Local Similarity   35.8%;  Pred. No. 1.5e-16;
  Matches   59;  Conservative   24;  Mismatches   71;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |||:
Db      87 FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFLGKDT 143

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |   |
Db     144 TKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATNQ 199

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           :    |   :  |||:      |:| ::    ||:   :||:  |
Db     200 IPTTLGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 15
US-09-791-537-79889
; Sequence 79889, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 79889
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-791-537-79889

  Query Match             18.4%;  Score 245.5;  DB 21;  Length 242;
  Best Local Similarity   35.8%;  Pred. No. 1.5e-16;
  Matches   59;  Conservative   24;  Mismatches   71;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |||:
Db      87 FDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFLGKDT 143

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |   |
Db     144 TKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATNQ 199

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           :    |   :  |||:      |:| ::    ||:   :||:  |
Db     200 IPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241
```

```
          :| ::  :|:   |             ||:| :      |:| : ||     ::    :
Db   1282 AGVDIQLMEDFVLGVSGAAFLGKMDSQKFDAEVSRKGVVGSVYTGFLAGSWFFKGQYSLG 1341

Qy     50 DGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNY 109
          : :  ::    |||::  |: |  :     |    |    | |         | | : | |
Db   1342 ETQNDMKTRYGVLGESSASWTSRGVL----ADALVEYRSLVGPVRPT---FYALHFNP-Y 1393

Qy    110 VR----------WTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPL 159
          |           :|  |:| |  |   ::|:|   |::| |     |   ||| |
Db   1394 VEVSYASMKFPGFTEQGREARSFEDASLTNITIPLGMKFEL-----AFIKGQFSE----- 1443

Qy    160 FQFIN-------WVKVYKYTPGQG----EGGSDF----------------TLDWTDNF 190
             :|       | : |:   |      | | |:                   :|:  |
Db   1444 ---VNSLGISYAW-EAYRKVEGGAVQLLEAGFDWEGAPMDLPRQELRVALENNTEWSSYF 1499

Qy    191 DTFDG 195
            |  |
Db   1500 STVLG 1504
```

Search completed: January 9, 2003, 12:14:40
Job time : 7.79861 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:57 ; Search time 27.8125 Seconds
(without alignments)
1279.206 Million cell updates/sec

Title: US-09-654-652A-2
Perfect score: 1439
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....PNSSSVDKLAAALEHHHHHH 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	257.5	17.9	308	20	AAW93001	B. alkalophilus be
2	246.5	17.1	242	10	AAP95000	Bacillus subtilis
3	244	17.0	214	22	AAE07317	Barley recombinant
4	244	17.0	239	11	AAR06621	Hybrid (1,3-1,4)-p
5	226	15.7	234	11	AAR03775	Thermostable beta-
6	226	15.7	237	11	AAR05803	Heat-stable endo-b
7	225	15.6	208	23	ABB76858	Bacterial lichenas
8	225	15.6	214	23	ABB76859	Bacterial lichenas
9	220.5	15.3	237	11	AAR06622	Hybrid (1,3-1,4)-p
10	189	13.1	245	19	AAW37884	Lichenase protein.
11	184	12.8	282	21	AAG48467	Arabidopsis thalia
12	184	12.8	282	23	ABB93737	Herbicidally activ
13	172	12.0	282	21	AAG32464	Arabidopsis thalia
14	172	12.0	282	23	ABB91456	Herbicidally activ
15	171.5	11.9	277	21	AAG52114	Arabidopsis thalia
16	171.5	11.9	277	23	ABB93169	Herbicidally activ
17	169.5	11.8	269	21	AAG18644	Arabidopsis thalia
18	169.5	11.8	282	21	AAG18643	Arabidopsis thalia
19	169.5	11.8	282	23	ABB93168	Herbicidally activ
20	169.5	11.8	298	21	AAG18642	Arabidopsis thalia
21	163	11.3	286	23	ABB93103	Herbicidally activ
22	161	11.2	284	23	ABB93853	Herbicidally activ
23	159	11.0	287	23	ABB93104	Herbicidally activ
24	154.5	10.7	284	23	ABB93854	Herbicidally activ
25	154	10.7	845	22	AAB98062	Clostridium perfri
26	150.5	10.5	269	21	AAG20347	Arabidopsis thalia
27	150.5	10.5	269	21	AAG20350	Arabidopsis thalia
28	150.5	10.5	269	21	AAG52653	Arabidopsis thalia
29	150.5	10.5	269	23	ABB93167	Herbicidally activ
30	150.5	10.5	290	21	AAG20346	Arabidopsis thalia
31	150.5	10.5	297	21	AAG52652	Arabidopsis thalia
32	150.5	10.5	300	21	AAG20349	Arabidopsis thalia
33	149.5	10.4	269	21	AAG20329	Arabidopsis thalia
34	149.5	10.4	299	21	AAG20328	Arabidopsis thalia
35	149.5	10.4	306	17	AAR97362	Oerskovia beta-1,3
36	147	10.2	247	18	AAW11593	Protein encoded by
37	147	10.2	247	20	AAY08308	A. thaliana merist
38	145.5	10.1	285	23	ABB93852	Herbicidally activ
39	145	10.1	263	18	AAW29455	Oerskovia xanthine
40	143	9.9	204	21	AAG48468	Arabidopsis thalia
41	143	9.9	435	18	AAW29456	Oerskovia xanthine
42	141	9.8	261	12	AAR11599	Beta-1,3-glucanase
43	140.5	9.8	303	18	AAW29457	Oerskovia xanthine
44	140.5	9.8	307	21	AAG40361	Arabidopsis thalia
45	140.5	9.8	310	21	AAG40360	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW93001
ID AAW93001 standard; Protein; 308 AA.
XX
AC AAW93001;
XX
DT 19-MAY-1999 (first entry)
XX
DE B. alkalophilus beta-glucanase protein.
XX
KW Beta-glucanase; enzyme; glucan; removal; lichenan; membrane;
KW equipment; food industry; brewing.
XX
OS Bacillus alkalophilus.
XX
PN DE19732751-A1.
XX
PD 04-FEB-1999.
XX
PF 30-JUL-1997; 97DE-1032751.
XX
PR 30-JUL-1997; 97DE-1032751.
XX
PA (HENK) HENKEL KGAA.
XX
PI Hillen W, Maurer K;
XX
DR WPI; 1999-122161/11.
DR N-PSDB; AAX02912.
XX
PT New Bacillus alkalophilus beta-glucanase protein and gene - useful
PT for removing glucan and/or lichenan from membranes in the brewing
PT industry

```
XX
PS   Claim 6; Page 5; 8pp; German.
XX
CC   This sequence represents a novel beta-glucanase isolated from Bacillus
CC   alkalophilus DSM 9956. This enzyme is useful for removing glucan and/or
CC   lichenan from membranes and equipment in the food industry, especially
CC   the brewing industry.
XX
SQ   Sequence   308 AA;

  Query Match             17.9%;  Score 257.5;  DB 20;  Length 308;
  Best Local Similarity   36.5%;  Pred. No. 1.4e-16;
  Matches   62;  Conservative   22;  Mismatches   73;  Indels   13;  Gaps    6;

Qy       6 DFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN 65
           ::   || | :  ||| ||  || | ::||||||:| |  |    |  || |:||| |||:
Db      96 EYKAGELRTNQFYQYGLFEVNMKPAKSTGTVSSLFTY-TGPWDWDNDPWDEIDIEFLGKD 154

Qy      66 PGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGG 125
             | |  |   |    :| :| :   | ::|:||  || |  : | |:|: |
Db     155 TTRVQFNYFTNGVG---NNEHYHELGFDASESFNTYAFEWRPESISWYVNGELVYTA--- 208

Qy     126 QVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             |:  | | :  |||       | | ||    |:    :||   :|||
Db     209 -TENIPQTPQKIMMNLWPGIGVDGWTGVFDGEDTPVVTEYDWV---RYTP 254


RESULT 2
AAP95000
ID   AAP95000 standard; Protein; 242 AA.
XX
AC   AAP95000;
XX
DT   12-FEB-1997  (first entry)
XX
DE   Bacillus subtilis lichenase.
XX
KW   Beer production; fermentation; barley; beta-glucan; hydrolysis;
KW   lichenase.
XX
OS   Bacillus subtilis Y-25.
XX
PN   JP01067181-A.
XX
PD   13-MAR-1989.
XX
PF   08-SEP-1987;   87JP-0224615.
XX
PR   08-SEP-1987;   87JP-0224615.
XX
PA   (ASAK ) ASAHI BREWERIES KK.
XX
DR   WPI; 1989-119863/16.
DR   N-PSDB; AAN95000.
XX
PT   Recombinant plasmid used in beer prodn. - obtd. by integrating
PT   lichenase gene derived from Bacillus subtilis, into vector
XX
PS   Disclosure; Fig 4; 7pp; Japanese.
XX
CC   The lichenase gene from Bacillus subtilis Y-25 is used for
CC   transforming Bacillus hosts so that they show increased lichenase
CC   expression. The recombinant lichenase enzyme produced by the
CC   transformants is useful in beer production for decomposing beta-
CC   glucan from barley.
XX
SQ   Sequence   242 AA;

  Query Match             17.1%;  Score 246.5;  DB 10;  Length 242;
  Best Local Similarity   35.5%;  Pred. No. 1.2e-15;
  Matches   60;  Conservative   24;  Mismatches   74;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |
Db      83 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFL 139

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
Db     140 GKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 195

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
             |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     196 ATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 3
AAE07317
ID   AAE07317 standard; Protein; 214 AA.
XX
AC   AAE07317;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Barley recombinant, thermostable (1,3-1,4)-beta-glucanase enzyme.
XX
KW   Barley; transgenic barley malt; carbohydrate degrading enzyme; chicken;
KW   (1,3-1,4)-beta-glucanase; nutritional value; foodstuff; thermostable.
XX
OS   Hordeum vulgare.
XX
PN   WO200159141-A2.
XX
PD   16-AUG-2001.
XX
PF   09-FEB-2001; 2001WO-US04222.
XX
PR   10-FEB-2000; 2000US-0181473.
PR   09-NOV-2000; 2000US-0247126.
XX
PA   (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Von Wettstein D,  Huang J,  Horvath H;
XX
DR   WPI; 2001-497082/54.
XX
PT   New barley based foodstuff for animals, i.e. chicken comprising
PT   recombinant carbohydrate degrading enzyme (1,3-1,4)-beta-glucanase -
XX
PS   Claim 8; Page 37-38; 43pp; English.
XX
CC   The present invention relates to a food stuff comprising barley feed and
CC   transgenic barley malt where the transgenic barley malt comprises a
CC   recombinant carbohydrate degrading enzyme comprising a (1,3-1,4)-beta-
CC   glucanase. The food stuff is useful to feed animal, preferably chickens
CC   Barley is cheaper than corn, the principal foodstuff of chickens.
CC   Chickens cannot efficiently utilise barley as an energy source as they
CC   do not possess a gut enzyme that depolymerises beta-D-glucan, major
CC   carbohydrate present in the barley endosperm. The invention provides a
CC   barley based foodstuff which comprises a recombinant carbohydrate
CC   degrading enzyme improving the nutritional value of the foodstuff.
CC   The present sequence is barley recombinant, thermostable
CC   (1,3-1,4)-beta-glucanase enzyme.
XX
SQ   Sequence   214 AA;

  Query Match             17.0%;  Score 244;  DB 22;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 1.7e-15;
  Matches   61;  Conservative   26;  Mismatches   69;  Indels   16;  Gaps    8;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      55 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 111

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db     112 GKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT 167
```

```
Qy    123 EGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
              :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db    168 ---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 212
```

RESULT 4
AAR06621
ID AAR06621 standard; protein; 239 AA.
XX
AC AAR06621;
XX
DT 09-JAN-1991 (first entry)
XX
DE Hybrid (1,3-1,4)-pre-beta-glucanase.
XX
KW Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.
XX
OS Bacillus amyloliquefaciens, Bacillus macerans.
XX
FH Key Location/Qualifiers
FT Domain 1..129
FT /label=amino terminal of beta-amyloliquefaciens
FT Domain 133..236
FT /label=carboxyl-terminal of B.macerans
XX
PN WO9009436-A.
XX
PD 23-AUG-1990.
XX
PF 16-FEB-1990; 90WO-DK00044.
XX
PR 04-AUG-1989; 89DK-0003848.
PR 16-FEB-1989; 89DD-0325800.
XX
PA (CARL-) CARLSBERG A/S.
PA (DEAK) AKAD WISSENSCHAFT DDR.
XX
PI Borriss R, Hofemeister J, Thomsen KK, Olsen O, Vonwettstein D;
XX
DR WPI; 1990-275129/36.
DR N-PSDB; AAQ05832.
XX
PT New thermostable (1,3-1,4)-beta-glucanase - prepd. using hybrid
PT gene obtd. using Bacillus amyloliquefaciens and B.macerans genes
XX
PS Disclosure; page 26; 84pp; English.
XX
CC This hybrid protein is encoded by the beta-glucanase-H1 gene.
CC Following processing of the signal peptide the mature protein
CC is produced, comprising the amino terminus of the amylolique-
CC faciens beta-glucanase and the carboxyl-terminal half of the
CC B.macerans beta-glucanase. This hybrid protein is thermostable
CC and hydrolyses beta-glycosidic linkages in (1,3-1,4)-beta-glucans.
CC Reducing sugars are obtd. at high temps. and thus this enzyme can
CC be used in the mfr. of food prods., esp. beer and animal feed (eg
CC for feeding poultry). See also AAQ05833.
XX
SQ Sequence 239 AA;

```
  Query Match             17.0%;  Score 244;  DB 11;  Length 239;
  Best Local Similarity   35.5%;  Pred. No. 2e-15;
  Matches   61;  Conservative   26;  Mismatches   69;  Indels   16;  Gaps    8;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db     80 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 136

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
          ||:   | |  |   |    ||  ::   | : ||||  :| | |::| |||  :: |
Db    137 GKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT 192

Qy    123 EGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
              :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db    193 ---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 237
```

RESULT 5
AAR03775
ID AAR03775 standard; protein; 234 AA.
XX
AC AAR03775;
XX
DT 31-JUL-1990 (first entry)
XX
DE Thermostable beta-glucanase.
XX
KW Thermostable beta-glucanase; ss; Bacillus subtilis.
XX
OS Bacillus macerans.
XX
PN DD272102-A.
XX
PD 27-SEP-1989.
XX
PF 12-MAY-1988; 88DD-0315706.
XX
PR 12-MAY-1988; 88DD-0315706.
XX
PA (DEAK) AKAD WISSENSCHAFT DDR.
XX
PI Borriss R;
XX
DR WPI; 1990-067913/10.
DR N-PSDB; AAQ03519.
XX
PT Thermostable beta-glucanase production -
PT using Bacillus subtilis transformed with gene from Bacillus
PT macerans.
XX
PS Fig 1; ; 9pp; German.
XX
CC The gene encoding Bacillus macerans thermostable
CC beta-glucanase is expressed in Bacillus subtilis. The
CC enzyme is useful for lowering the viscosity of brewing
CC mashes and in the production of feedstuff.
XX
SQ Sequence 234 AA;

```
  Query Match             15.7%;  Score 226;  DB 11;  Length 234;
  Best Local Similarity   35.1%;  Pred. No. 1.1e-13;
  Matches   59;  Conservative   23;  Mismatches   70;  Indels   16;  Gaps    8;

Qy      7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
          |  ||  :     || :|  || |  :| ||| | |      | |  | |:||| |||:
Db     79 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 135

Qy     67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
            | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db    136 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 188

Qy    127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
              :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db    189 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 232
```

RESULT 6
AAR05803
ID AAR05803 standard; protein; 237 AA.
XX
AC AAR05803;
XX
DT 08-NOV-1990 (first entry)
XX
DE Heat-stable endo-beta-1,3-1,4-glucanase.
XX

KW Endo-beta-1,3-1,4-glucanase; barley; brewing.
XX
OS Bacillus macerans.
XX
PN DD275704-A.
XX
PD 31-JAN-1990.
XX
PF 23-SEP-1988; 88DD-0320082.
XX
PR 23-SEP-1988; 88DD-0320082.
XX
PA (DEAK) AKAD WISSENSCHAFT DDR.
XX
PI Borriss R, Wobus U, Mendel R-R, Baumlein H;
XX
DR WPI; 1990-210631/28.
DR N-PSDB; AAQ05167.
XX
PT Prepn. of barley plants expressing heat stable beta-glucanase -
PT by transforming cells with appropriate vector then regeneration
PT giving seeds useful in brewing without conversion to malt.
XX
PS Disclosure; ; p; German.
XX
CC The corresp. gene and the plant promoter region are inserted into an
CC expression vector and used to transform barley cells. The transformants
CC are used to regenerate barley plants which are useful in brewing. This
CC protein accumulates in the ripe (but ungerminated) seeds. The
CC proportion of these to malted seeds is increased therefore, without
CC an unacceptable increase in viscosity.
XX
SQ Sequence 237 AA;

Query Match 15.7%; Score 226; DB 11; Length 237;
Best Local Similarity 35.1%; Pred. No. 1.1e-13;
Matches 59; Conservative 23; Mismatches 70; Indels 16; Gaps 8;

```
Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db      82 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
Db     139 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 191

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
Db     192 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 235
```

RESULT 7
ABB76858
ID ABB76858 standard; Protein; 208 AA.
XX
AC ABB76858;
XX
DT 27-JUN-2002 (first entry)
XX
DE Bacterial lichenase #1.
XX
KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.
XX
OS Unidentified.
XX
PN WO200212511-A1.
XX
PD 14-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-ES00303.
XX
PR 28-JUL-2000; 2000ES-0001922.
XX
PA (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX
PI Rodriguez Munoz V, Perez Mellado R;
XX
DR WPI; 2002-217195/27.
DR N-PSDB; ABL53374.
XX
PT Producing nucleic acid encoding lichenase, useful for producing enzymes
PT for improving filtration in brewing, comprises isolating the nucleic
PT acid from soil bacteria -
XX
PS Claim 18; Page 18-19; 27pp; Spanish.
XX
CC The present sequence is a protein sequence from a soil bacterium for an
CC enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)
CC activity. The enzyme is useful in brewing and for degrading
CC beta-glucanases that cause problems during filtration.
XX
SQ Sequence 208 AA;

Query Match 15.6%; Score 225; DB 23; Length 208;
Best Local Similarity 36.2%; Pred. No. 1.2e-13;
Matches 55; Conservative 21; Mismatches 66; Indels 10; Gaps 5;

```
Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db      58 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 114

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
Db     115 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 170

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 153
Db     171 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYN 200
```

RESULT 8
ABB76859
ID ABB76859 standard; Protein; 214 AA.
XX
AC ABB76859;
XX
DT 27-JUN-2002 (first entry)
XX
DE Bacterial lichenase #2.
XX
KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.
XX
OS Unidentified.
XX
PN WO200212511-A1.
XX
PD 14-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-ES00303.
XX
PR 28-JUL-2000; 2000ES-0001922.
XX
PA (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX
PI Rodriguez Munoz V, Perez Mellado R;
XX
DR WPI; 2002-217195/27.
DR N-PSDB; ABL53375.
XX
PT Producing nucleic acid encoding lichenase, useful for producing enzymes
PT for improving filtration in brewing, comprises isolating the nucleic
PT acid from soil bacteria -
XX
PS Claim 18; Page 21-22; 27pp; Spanish.
XX
CC The present sequence is a protein sequence from a soil bacterium for an

```
CC   enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)
CC   activity. The enzyme is useful in brewing and for degrading
CC   beta-glucanases that cause problems during filtration.
XX
SQ   Sequence   214 AA;

  Query Match             15.6%;  Score 225;  DB 23;  Length 214;
  Best Local Similarity   36.2%;  Pred. No. 1.2e-13;
  Matches   55;  Conservative   21;  Mismatches   66;  Indels   10;  Gaps    5;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db      63 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 119

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
Db     120 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 175

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 153
Db     176 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYN 205


RESULT 9
AAR06622
ID   AAR06622 standard; protein; 237 AA.
XX
AC   AAR06622;
XX
DT   09-JAN-1991  (first entry)
XX
DE   Hybrid (1,3-1,4)-pre-beta-glucanase.
XX
KW   Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.
XX
OS   Bacillus amyloliquefaciens, Bacillus macerans.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..127
FT                   /label=amino-terminal half of B.macerans beta-glucanase
FT   Domain          131..234
FT                   /label=carboxyl-terminal half of B.amyloliquefaciens
FT                         beta-glucanase
XX
PN   WO9009436-A.
XX
PD   23-AUG-1990.
XX
PF   16-FEB-1990;   90WO-DK00044.
XX
PR   04-AUG-1989;   89DK-0003848.
PR   16-FEB-1989;   89DD-0325800.
XX
PA   (CARL-) CARLSBERG A/S.
PA   (DEAK ) AKAD WISSENSCHAFT DDR.
XX
PI   Borriss R,  Hofemeister J,  Thomsen KK,  Olsen O,  Vonwettstein D;
XX
DR   WPI; 1990-275129/36.
DR   N-PSDB; AAQ05833.
XX
PT   New thermostable (1,3-1,4)-beta-glucanase - prepd. using hybrid
PT   gene obtd. using Bacillus amyloliquefaciens and B.macerans genes
XX
PS   Disclosure; page 28; 84pp; English.
XX
CC   This hybrid protein is encoded by the beta-glucanase-H2 gene.
CC   Following processing of the signal peptide the mature protein
CC   is produced, comprising the amino terminus of the B.macerans
CC   beta-glucanase and the carboxyl-terminal half of the B.amylo-
CC   liquefaciens beta-glucanase.  This hybrid protein is thermo-
CC   stable and hydrolyses beta-glycosidic linkages in (1,3-1,4)-
CC   beta-glucans.  Reducing sugars are obtd. at high temps. and
CC   thus this enzyme can be used in the mfr. of food prods., esp.
CC   beer and animal feed (e.g.for feeding poultry).  See also
CC   AAQ05833.
XX
SQ   Sequence   237 AA;

  Query Match             15.3%;  Score 220.5;  DB 11;  Length 237;
  Best Local Similarity   33.9%;  Pred. No. 3.9e-13;
  Matches   56;  Conservative   23;  Mismatches   75;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db      82 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
Db     139 TKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATTQ 194

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
Db     195 IPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYHWMRYRK 236


RESULT 10
AAW37884
ID   AAW37884 standard; Protein; 245 AA.
XX
AC   AAW37884;
XX
DT   20-AUG-1998  (first entry)
XX
DE   Lichenase protein.
XX
KW   Lichenase; licA; fungus; enzyme; beta-1,4-glucan bond hydrolysation;
KW   beta-1,3-linked glucan bond; grain-containing feed; grain treatment;
KW   nutrient availability; brewing; fermentation.
XX
OS   Orpinomyces sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..29
FT                   /note= "signal peptide"
FT   Protein         30..246
FT                   /note= "mature lichenase"
XX
PN   WO9814595-A1.
XX
PD   09-APR-1998.
XX
PF   03-OCT-1997;   97WO-US17811.
XX
PR   04-OCT-1996;   96US-0027882.
XX
PA   (UYGE-) UNIV GEORGIA RES FOUND INC.
XX
PI   Chen H,  Li X,  Ljungdahl LG;
XX
DR   WPI; 1998-240094/21.
DR   N-PSDB; AAV29067.
XX
PT   New isolated lichenase protein - is obtained from Orpinomyces PC-2,
PT   used for treatment of grain to improve feeds or to improve brewing
PT   and fermentation processes
XX
PS   Claim 1; Page 24-25; 41pp; English.
XX
CC   This sequence is the lichenase (licA) of Orpinomyces sp. strain PC-2
CC   of the invention. The protein was purified from a fungus or a fungal
CC   culture or from a recombinant DNA molecule having a fungal lichenase
CC   coding sequence. The lichenase enzyme hydrolyses the beta-1,4-glucan
CC   bonds adjacent to beta-1,3-linked glucan bonds, but does not cleave
CC   beta-1,4-linked glucans. They can be used for the treatment of animal
CC   grain-containing feeds to improve nutrient availability and for treatment
```

```
CC   of grain (e.g. barley or wheat) in the brewing and fermentation
CC   industries to increase carbon substrate availability and to maximise
CC   production of desired products.
XX
SQ   Sequence   245 AA;

  Query Match             13.1%;  Score 189;  DB 19;  Length 245;
  Best Local Similarity   32.9%;  Pred. No. 4.7e-10;
  Matches   56;  Conservative   16;  Mismatches   76;  Indels   22;  Gaps    8;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           :| ||   ||  | :      || |:  ||     | ||| | |   |   ||  | |:|
Db      83 RDGSG---YTCGEYRTKNYYGYGMFQVNMKPIKNPGVVSSFFTYTGPS---DGTKWDEID 136

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || || :    | |  |    |     |  | :   | | |||||  |  | : | |||
Db     137 IEFLGYDTTKVQFNYYTNGQGHH---EHIHYLGFDASQGFHTYGFFWARNSITWYVDGTA 193

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWV 166
           |         |:  | | :  | |:      |:  |:  :  :  : :||
Db     194 VYTA----YDNIPDTPGKIMMNAWNGIGVDDWLRPFN-GRTNISAYYDWV 238


RESULT 11
AAG48467
ID   AAG48467 standard; Protein; 282 AA.
XX
AC   AAG48467;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 61207.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 12.8%; Score 184; DB 21; Length 282;
Best Local Similarity 27.6%; Pred. No. 1.8e-09;
Matches 59; Conservative 33; Mismatches 98; Indels 24; Gaps 9;

```
Qy       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDI 59
           :|   |||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:|
Db      54 LSLDKFSGSGFQSHQEFLYGKVEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDF 107

Qy      60 EVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 116
           | ||    :| :  :|: | |    |  : |    |  |  |||| : | |  | :|:||
Db     108 EFLGNISGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVD--FHTYCIIWNPQRVIFTIDG 164

Qy     117 QEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKVY 169
             :|: :  :   :     | :|   :|| :|  |  |   : | || |   |     |
Db     165 IPIREFKNSEALGVPFPKHQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAFYRNYNVD 224

Qy     170 KYTPGQGEGGSDFTLDWTDNFDTFDGS---RWGK 200
                 |:        |      | |    :| :
Db     225 ACVWSNGKSSCSANSSWFTQVLDFKGKNRVKWAQ 258
```

RESULT 12
ABB93737
ID ABB93737 standard; Protein; 282 AA.
XX
AC ABB93737;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2948.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 2948; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX

```
SQ   Sequence   282 AA;

  Query Match             12.8%;  Score 184;  DB 23;  Length 282;
  Best Local Similarity   27.6%;  Pred. No. 1.8e-09;
  Matches   59;  Conservative   33;  Mismatches   98;  Indels   24;  Gaps    9;

Qy       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDI 59
           :|   |||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:|
Db      54 LSLDKFSGSGFQSHQEFLYGKVEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDF 107

Qy      60 EVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 116
           | ||    :| :  :|: | |    |  : |    |  |  ||||.: | |  | :|:||
Db     108 EFLGNISGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVD--FHTYCIIWNPQRVIFTIDG 164

Qy     117 QEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKVY 169
             :|: :  :   :     | :|   :|| :|  |  |   : | || |   |     |
Db     165 IPIREFKNSEALGVPFPKHQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAFYRNYNVD 224

Qy     170 KYTPGQGEGGSDFTLDWTDNFDTFDGS---RWGK 200
               |:        |      | |    :| :
Db     225 ACVWSNGKSSCSANSSWFTQVLDFKGKNRVKWAQ 258


RESULT 13
AAG32464
ID   AAG32464 standard; Protein; 282 AA.
XX
AC   AAG32464;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 39167.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 12.0%; Score 172; DB 21; Length 282;
Best Local Similarity 25.9%; Pred. No. 2.6e-08;
Matches 63; Conservative 34; Mismatches 88; Indels 58; Gaps 12;

```
Qy       8 SGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK- 64
           ||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:| | ||
Db      60 SGSGFQSNQEFLYGKAEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDFEFLGNI 113

Qy      65 --NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
             :| :  :|: | |    |  : |    |  :  |||| : | |  : :|||  :|:
Db     114 SGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVN--FHTYCITWNPQRIIFTVDGIPIREF 170

Qy     123 EGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKVYKYTPGQ 175
           :  :   :     | :|   :|| :|  |  |   : | || |   |      |:
Db     171 KNPEAIGVPFPTRQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAF------YR----- 219

Qy     176 GEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQES 235
                       |::  ||  |  |  :   |    | |           |   ||
Db     220 ------------NYNV-DGCVWANGKSSCSANSPWFTQK------------LDSNGQTR 253

Qy     236 FNG 238
             |
Db     254 MKG 256
```

RESULT 14
ABB91456
ID ABB91456 standard; Protein; 282 AA.
XX
AC ABB91456;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 667.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 667; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant

```
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   282 AA;

  Query Match            12.0%;  Score 172;  DB 23;  Length 282;
  Best Local Similarity   25.9%;  Pred. No. 2.6e-08;
  Matches   63;  Conservative   34;  Mismatches   88;  Indels   58;  Gaps   12;

Qy       8 SGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK- 64
           ||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:| | ||
Db      60 SGSGFQSNQEFLYGKAEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDFEFLGNI 113

Qy      65 --NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
             :| :  :|: | |    |  : |    |  :  |||| : | |  : :|||| :|:
Db     114 SGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVN--FHTYCITWNPQRIIFTVDGIPIREF 170

Qy     123 EGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKVYKYTPGQ 175
           :  :   :     | :|   :|| :|  |  |    : | || |   |      |:
Db     171 KNPEAIGVPFPTRQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAF------YR----- 219

Qy     176 GEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQES 235
                       |::  ||  |  |  :   |    | |           |   ||
Db     220 ------------NYNV-DGCVWANGKSSCSANSPWFTQK-----------LDSNGQTR 253

Qy     236 FNG 238
             |
Db     254 MKG 256


RESULT 15
AAG52114
ID   AAG52114 standard; Protein; 277 AA.
XX
AC   AAG52114;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 66212.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
```

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 11.9%; Score 171.5; DB 21; Length 277;
Best Local Similarity 27.0%; Pred. No. 2.8e-08;
Matches 53; Conservative 32; Mismatches 70; Indels 41; Gaps 9;

```
Qy       8 SGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK- 64
           ||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:| | ||
Db      55 SGSGFQSNQEFLYGKAEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDFEFLGNI 108

Qy      65 --NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
             :| :  :|: | |    |  : |    | |:  |||| : | |  : :|||| :|:
Db     109 SGHPYTLHTNVYT-KGSGDKEQQFHLWFDPTAN--FHTYCITWNPQRIIFTVDGIPIREF 165

Qy     123 EGGQVSNLTGTQGLRF----------NLWSSESAAWVG---QFDESKLPLFQFIN----- 164
             :      ::|: |          :|| :|  |  |   : | || |   :
Db     166 MNAE------SRGVPFPTKQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAYYRNYNVE 219

Qy     165 ---WVKVYKYTPGQGE 177
               ||      |   :
Db     220 GCVWVNGKSVCPANSQ 235
```

Search completed: January 9, 2003, 12:16:18
Job time : 28.8125 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:13:37 ; Search time 22.1019 Seconds
(without alignments)
2312.009 Million cell updates/sec

Title: US-09-654-652A-1
Perfect score: 1333
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....TRKGQESFNGQVPRDDEPAP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	266.5	20.0	851	16	Q9K7X5	Q9k7x5 bacillus ha
2	248.5	18.6	256	2	Q9APD8	Q9apd8 bacillus ci
3	247.5	18.6	276	2	Q45648	Q45648 bacillus sp
4	242.5	18.2	214	2	Q93GE8	Q93ge8 uncultured
5	242.5	18.2	802	2	Q9S310	Q9s310 ruminococcu
6	241.5	18.1	214	2	Q93GE7	Q93ge7 uncultured
7	235.5	17.7	214	2	Q93GE6	Q93ge6 uncultured
8	234.5	17.6	239	2	Q45691	Q45691 bacillus su
9	232.5	17.4	242	2	Q8RMP0	Q8rmp0 bacillus su
10	230	17.3	212	2	Q9FDC9	Q9fdc9 paenibacill
11	230	17.3	237	2	O07856	O07856 streptococc
12	221	16.6	246	16	Q97FD3	Q97fd3 clostridium
13	219	16.4	205	2	Q93UI1	Q93ui1 uncultured
14	206	15.5	289	16	Q8U8N5	Q8u8n5 agrobacteri
15	200	15.0	293	16	Q98C78	Q98c78 rhizobium l
16	190.5	14.3	302	16	Q9K7X6	Q9k7x6 bacillus ha
17	189	14.2	245	3	O14412	O14412 orpinomyces
18	184	13.8	282	10	Q9FI31	Q9fi31 arabidopsis
19	172	12.9	282	10	O80803	O80803 arabidopsis
20	171.5	12.9	277	10	Q9M0D1	Q9m0d1 arabidopsis
21	169.5	12.7	282	10	Q9M0D2	Q9m0d2 arabidopsis
22	169	12.7	642	16	Q9WXN1	Q9wxn1 thermotoga
23	168.5	12.6	277	10	Q94A49	Q94a49 arabidopsis
24	167.5	12.6	302	3	O42800	O42800 aspergillus
25	167	12.5	646	2	Q60039	Q60039 thermotoga
26	163	12.2	286	10	Q38910	Q38910 arabidopsis
27	161	12.1	284	10	Q9FKL8	Q9fkl8 arabidopsis
28	159	11.9	286	10	Q8W4M6	Q8w4m6 arabidopsis
29	159	11.9	287	10	Q9ZSU4	Q9zsu4 arabidopsis
30	157.5	11.8	92	3	Q9HGU1	Q9hgu1 aspergillus
31	157.5	11.8	94	3	Q9P420	Q9p420 trichoderma
32	156	11.7	292	10	Q9ZRV1	Q9zrv1 fagus sylva
33	154.5	11.6	277	10	Q38907	Q38907 arabidopsis
34	154.5	11.6	284	10	Q9SEB0	Q9seb0 arabidopsis
35	154.5	11.6	306	2	Q51333	Q51333 oerskovia x
36	154	11.6	845	2	Q9KWF3	Q9kwf3 clostridium
37	154	11.6	845	16	Q8XNF8	Q8xnf8 clostridium
38	150.5	11.3	269	10	Q39148	Q39148 arabidopsis
39	149	11.2	163	2	Q93UI2	Q93ui2 uncultured
40	146.5	11.0	284	10	Q38857	Q38857 arabidopsis
41	145.5	10.9	285	10	Q9FKL9	Q9fkl9 arabidopsis
42	143	10.7	435	2	O68641	O68641 oerskovia x
43	141.5	10.6	285	10	Q8S902	Q8s902 glycine max
44	141.5	10.6	289	10	P93669	P93669 hordeum vul
45	141	10.6	280	10	Q949I0	Q949i0 festuca pra

ALIGNMENTS

RESULT 1
Q9K7X5
ID Q9K7X5 PRELIMINARY; PRT; 851 AA.
AC Q9K7X5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Endo-beta-1,3-1,4 glucanase (licheninase) (EC 3.2.1.73).
GN BGLS OR BH3232.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001518; BAB06951.1; -.
DR HSSP; P23904; 1AJK.
DR InterPro; IPR000757; Glyco_hydro_16.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00722; Glyco_hydro_16; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR PRINTS; PR00737; GLHYDRLASE16.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 3.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 851 AA; 98142 MW; C1C4F0F158400285 CRC64;

Query Match 20.0%; Score 266.5; DB 16; Length 851;
Best Local Similarity 29.7%; Pred. No. 1.8e-13;
Matches 82; Conservative 31; Mismatches 108; Indels 55; Gaps 12;

Qy 4 AKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLG 63

```
Db     99 APPYKAGELRTNDFYQYGLFEVSMKPAKVEGTVSSFFTY-TGEWDWDGDPWDEIDIEFLG 157

Qy     64 KNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
Db    158 KDTTRIQFNYFTNGVGG---NEFYYDLGFDASESFNTYAFEWREDSITWYVNGEAVHTA- 213

Qy    124 GGQVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP----GQGE 177
Db    214 ---TENIPQTPQKIMMNLWPGVGVDEWTGVFDGDNTPLHADYEWV---RYTPLEALDEEE 267

Qy    178 GGSDFTLDWT----------------------DNFDTFDGSRWG-KGDWTFDGNRVDLT 213
Db    268 SGDDNEEEPVEEVEEEPAEDEEVSVRPIGSAIYETFNTFNEDIWSIAHGWT-NGQMFNAT 326

Qy    214 --DKNIYSRDGMLILALTRKGQESFNGQVPRDDEPA 247
Db    327 WYDSQITFSNGIMRFAIDKE----------DDEEA 351


RESULT 2
Q9APD8 .
ID   Q9APD8      PRELIMINARY;      PRT;   256 AA.
AC   Q9APD8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-beta-1,3-1,4-glucanase.
GN   BGLBC1.
OS   Bacillus circulans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1397;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC21367;
RA   Lee D.-S., Kim J.-Y., Kim H.-B.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF306531; AAG53947.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
SQ   SEQUENCE   256 AA;  27604 MW;  989E50C09F6E4531 CRC64;

  Query Match             18.6%;  Score 248.5;  DB 2;  Length 256;
  Best Local Similarity   34.7%;  Pred. No. 1.1e-12;
  Matches   61;  Conservative   28;  Mismatches   70;  Indels   17;  Gaps    7;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db     91 SGKPYASAEYATTQKYGYGRVEARIKAAKGTGLVTSLFTY---SGAAPGTSNDEIDIEIL 147

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQA--FHTYGLEWTPNYVRWTVDGQEVR 120
Db    148 GKDTTKMETNYFTNGVGGHST-----VIDLGFDASLDFHDYAFEWSPTSIKWYVDGRLVH 202

Qy    121 KTEGGQVSNLTGTQG-LRFNLWSSESAA--WVGQFDESKLPLFQFINWVKVYKYTP 173
Db    203 -TETGSRGPLPTSPGYIMVNLWSGAGPAEIWTGKFTYPGHPIRAYYDWI---KFTP 254


RESULT 3
Q45648
ID   Q45648      PRELIMINARY;      PRT;   276 AA.
AC   Q45648;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Lichenase precursor.
GN   BGA1.
OS   Bacillus sp.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1409;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=137;
RX   MEDLINE=94288605; PubMed=7517127;
RA   Tabernero C., Coll P.M., Fernandez-Abalos J.M., Perez P.,
RA   Santamaria R.I.;
RT   "Cloning and DNA sequencing of bgaA, a gene encoding an endo-beta-1,3-
RT   1, 4-glucanase, from an alkalophilic Bacillus strain (N137).";
RL   Appl. Environ. Microbiol. 60:1213-1220(1994).
DR   EMBL; Z12151; CAA78135.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Signal.
FT   SIGNAL        1     31       POTENTIAL.
FT   CHAIN        32    276       LICHENASE.
SQ   SEQUENCE   276 AA;  31770 MW;  D047F8A34CA9EBE2 CRC64;

  Query Match             18.6%;  Score 247.5;  DB 2;  Length 276;
  Best Local Similarity   36.1%;  Pred. No. 1.5e-12;
  Matches   61;  Conservative   20;  Mismatches   75;  Indels   13;  Gaps    6;

Qy      7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db     96 YKAGELRTNDFYHYGLFEVSMKPAKSTGTVSSFFTY-TGPWDWENDPWDEIDIEFLGKDT 154

Qy     67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
Db    155 TKIQFNYFTNGVGG---NEHYHELGFDAADDFNTYAFEWRPESIRWFVNGELVHTA---- 207

Qy    127 VSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
Db    208 TENIPQTPQKIMMNLWPGIGVDGWTGRFNGEDTPVVTQYDWV---KYTP 253


RESULT 4
Q93GE8
ID   Q93GE8      PRELIMINARY;      PRT;   214 AA.
AC   Q93GE8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254957; AAK50610.2; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   214 AA;  24159 MW;  3DB186D3B9D991CF CRC64;

  Query Match             18.2%;  Score 242.5;  DB 2;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 2.7e-12;
  Matches   60;  Conservative   24;  Mismatches   74;  Indels   11;  Gaps    6;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db     55 SYNKFDCGENRSVQTYGYGPYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 111

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
```

```
Db    112 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 167

Qy    123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
          |:    |   :  |||:      |:| ::    ||:   :||:  |
Db    168 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 213


RESULT 5
Q9S310
ID   Q9S310      PRELIMINARY;      PRT;   802 AA.
AC   Q9S310;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Family 11 xylanase /family 16 beta (1,3-1,4) glucanase.
GN   XYND.
OS   Ruminococcus flavefaciens.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Lachnospiraceae; Ruminococcus.
OX   NCBI_TaxID=1265;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=17;
RA   Aurilia V., Martin J.C., Munro C.A., Mercer D.K., Flint H.J.;
RT   "Organisation and strain distribution of genes responsible for the
RT   utilization of xylans by the rumen cellulolytic bacterium Ruminococcus
RT   flavefaciens 17.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ132472; CAB51934.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR003305; CBM_CenC.
DR   InterPro; IPR001137; GH_11.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF02018; CBM_4_9; 2.
DR   Pfam; PF00457; Glyco_hydro_11; 2.
DR   Pfam; PF00722; Glyco_hydro_16; 2.
DR   PRINTS; PR00911; GLHYDRLASE11.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS00776; GLYCOSYL_HYDROL_F11_1; 2.
DR   PROSITE; PS00777; GLYCOSYL_HYDROL_F11_2; 2.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 2.
FT   CHAIN         2    802       FAMILY 11 XYLANASE /FAMILY 16 BETA
FT                                (1,3-1,4) GLUCANASE.
SQ   SEQUENCE   802 AA;  89019 MW;  F36BC68805FC5274 CRC64;

  Query Match             18.2%;  Score 242.5;  DB 2;  Length 802;
  Best Local Similarity   34.9%;  Pred. No. 1.6e-11;
  Matches   58;  Conservative   21;  Mismatches   74;  Indels   13;  Gaps    6;

Qy      7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
          :|| |  |     || :|  |:     | ||| | |   |   |  |! |:|||:|||
Db    634 YSGGEFRTNNFYHYGYYECSMQAMKNDGVVSSFFTYTGPS---DDNPWDEIDIEILGKNT 690

Qy     67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
            | |  |   |     || : :   : :|:|||| :| |||: | |||:|| :
Db    691 TQVQFNYYTNGQGKH---EKLYDLGFDSSEAYHTYGFDWQPNYIAWYVDGREVYRA---- 743

Qy    127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
            ::  | | :  | |   :   |:  |:  : ||     ||   |
Db    744 TQDIPKTPGKIMMNAWPGLTVDDWLKAFN-GRTPLTAHYQWVTYNK 788


RESULT 6
Q93GE7
ID   Q93GE7      PRELIMINARY;      PRT;   214 AA.
AC   Q93GE7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254959; AAK50612.2; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   214 AA;  24175 MW;  3DB192D2F6B4CAA3 CRC64;

  Query Match             18.1%;  Score 241.5;  DB 2;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 3.3e-12;
  Matches   60;  Conservative   24;  Mismatches   74;  Indels   11;  Gaps    6;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          |   |   | :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db     55 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 111

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
          ||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
Db    112 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 167

Qy    123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
            |:    |   :  |||:      |:| ::    ||:   :||:  |
Db    168 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 213


RESULT 7
Q93GE6
ID   Q93GE6      PRELIMINARY;      PRT;   214 AA.
AC   Q93GE6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254961; AAK50614.2; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   214 AA;  24131 MW;  1E9D8BFA4300EBA5 CRC64;

  Query Match             17.7%;  Score 235.5;  DB 2;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 1e-11;
  Matches   60;  Conservative   23;  Mismatches   75;  Indels   11;  Gaps    6;

Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          |   |   | :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db     55 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIESL 111

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
          ||:    | |  |  ||     ||   :   |  |:| |  :| || ::| |||| :: |
Db    112 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHMYAFDWQPNSIKWYVDGQ-LKHT 167

Qy    123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
            ||    |   :  |||:      |:| ::    ||:   :||:  |
Db    168 ATTQVPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PLYAHYDWVRYTK 213


RESULT 8
Q45691
ID   Q45691      PRELIMINARY;      PRT;   239 AA.
```

```
AC   Q45691;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-beta-1,3-1,4-glucanase.
GN   BEG1.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCIB 8565;
RA   van Rensburg P., van Zyl W.H., Pretorius I.S.;
RT   "Over-expression of the Saccharomyces cerevisiae exo-beta-1,3-
RT   glucanase gene together with the Butyrivibrio fibrisolvens endo-beta-
RT   1,4-glucanase gene and the Bacillus subtilis endo-beta-1,3-1,4-
RT   glucanase gene in Saccharomyces cerevisiae.";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U60830; AAB05759.1; -.
DR   HSSP; P27051; 1GBG.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
SQ   SEQUENCE   239 AA;  26969 MW;  C4219760D13F878A CRC64;

  Query Match             17.6%;  Score 234.5;  DB 2;  Length 239;
  Best Local Similarity   33.7%;  Pred. No. 1.4e-11;
  Matches   58;  Conservative   26;  Mismatches   71;  Indels   17;  Gaps    7;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      80 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 136

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPA---ADQAFHTYGLEWTPNYVRWTVDGQEV 119
           ||:    | |  |  |      | |  ::     |  |:||| :| || ::| || | :
Db     137 GKDTTKVQFNYYTNGA------ENHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDWQ-L 189

Qy     120 RKTEGGQVSNLTGTQGLRFNLWSSESAA-WVGQFDESKLPLFQFINWVKVYK 170
           : |   |:    |   :  |||:    : |:| ::    ||:   :||:  |
Db     190 KHTATTQIPTAPGK--IMMNLWNGTGVSDWLGSYNGVN-PLYAHYDWVRYTK 238


RESULT 9
Q8RMP0
ID   Q8RMP0      PRELIMINARY;      PRT;   242 AA.
AC   Q8RMP0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Beta-1,3-1,4-glucanase.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sun J., Li W., Gu S., Xu Z., Zhao H., Xiao J.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF490978; AAM08358.1; -.
SQ   SEQUENCE   242 AA;  27238 MW;  285FED3FE76AE69A CRC64;

  Query Match             17.4%;  Score 232.5;  DB 2;  Length 242;
  Best Local Similarity   34.9%;  Pred. No. 2.1e-11;
  Matches   59;  Conservative   24;  Mismatches   75;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |
Db      83 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFL 139

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
Db     140 GKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 195

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           |:    |   :  |||:       :| ::    ||:   :||:  |
Db     196 ATNQIPTTPGK--IMMNLWNGTGVDERLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 10
Q9FDC9
ID   Q9FDC9      PRELIMINARY;      PRT;   212 AA.
AC   Q9FDC9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-1,3-1,4-beta-glucanase (Fragment).
OS   Paenibacillus polymyxa (Bacillus polymyxa).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Paenibacillaceae; Paenibacillus.
OX   NCBI_TaxID=1406;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yao W., Wang Y., Song W., Yang K., Su Z.;
RT   "Gene cloning of an antifugal protein.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF284449; AAG02415.1; -.
DR   HSSP; P23904; 1AJK.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
FT   NON_TER       1      1
FT   NON_TER     212    212
SQ   SEQUENCE   212 AA;  24127 MW;  7E707E7FBEF7A440 CRC64;

  Query Match             17.3%;  Score 230;  DB 2;  Length 212;
  Best Local Similarity   35.1%;  Pred. No. 2.9e-11;
  Matches   59;  Conservative   23;  Mismatches   70;  Indels   16;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           | | |  :    :|| :|  || |  :| ||| | |       |:|  | |:||| |||:
Db      57 FDGGEYRSKNTYRYGLYEVNMKPAKNTGIVSSFFTYTGP---ANGTQWDEIDIEFLGKDT 113

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
           | |  |   |     ||   :   |   ||||  :| | |::| |||  :: |
Db     114 TKVQFNYYTNGIGGH---EKVVDLGFDASSGFHTYAFDWQPGYIKWYVDG-VLKHT---A 166

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
           :|:  | | :  |||:     :|:| ::    ||:   :||   |||
Db     167 TTNIPKTPGQIMMNLWNGTGVDSWLGPYNGVN-PLYAEYDWV---KYT 210


RESULT 11
O07856
ID   O07856      PRELIMINARY;      PRT;   237 AA.
AC   O07856;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Beta-(1,3-1,4)-glucanase precursor.
GN   LICHENASE.
OS   Streptococcus bovis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Streptococcus.
OX   NCBI_TaxID=1315;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JB1;
RA   Ekinci M., Flint H.J.;
RT   "Isolation and overexpression of a gene encoding an extracellular
RT   beta-(1,3-1,4)-glucanase from Streptococcus bovis JB1.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z92911; CAB07443.1; -.
DR   HSSP; P23904; 1AJK.
```

```
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    237       BETA-(1,3-1,4)-GLUCANASE.
SQ   SEQUENCE   237 AA;  26989 MW;  7DEF5BCE53790470 CRC64;

  Query Match             17.3%;  Score 230;  DB 2;  Length 237;
  Best Local Similarity   33.1%;  Pred. No. 3.3e-11;
  Matches   54;  Conservative   26;  Mismatches   69;  Indels   14;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           ::| |  : |   || |:  ||     | ||| | |   |   ||  | |:||| |||:
Db      81 YTGGEWRSKERFGYGLFQVNMKPIKNPGVVSSFFTYTGPS---DGTKWDEIDIEFLGKDT 137

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |     |   |  : :   | | ||||| :|  ::: | |||: |
Db     138 TKVQFNYYT---SGQGNHEYLYNLGFDASQGFHTYGFDWQADHITWYVDGRAVYTA---- 190

Qy     127 VSNLTGTQG-LRFNLW--SSESAAWVGQFDESKLPLFQFINWV 166
           :|:  | | :  | |  : |  :|:| ::  : ||: : :|:
Db     191 YNNIPSTPGKIMMNAWPGTHEVDSWLGAYN-GRTPLYAYYDWI 232


RESULT 12
Q97FD3
ID   Q97FD3      PRELIMINARY;      PRT;   246 AA.
AC   Q97FD3;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Endo-1,3(4)-beta-glucanase family 16.
GN   CAC2807.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
DR   EMBL; AE007778; AAK80751.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW   Complete proteome.
SQ   SEQUENCE   246 AA;  27717 MW;  C0EBC302678D2FE1 CRC64;

  Query Match             16.6%;  Score 221;  DB 16;  Length 246;
  Best Local Similarity   33.7%;  Pred. No. 1.9e-10;
  Matches   58;  Conservative   19;  Mismatches   73;  Indels   22;  Gaps    8;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           ::| |  :     || :   || |   |  || | |   |   |  || |:||| |||:
Db      91 YAGGEYRSNNRYGYGLYRVSMKPAKHIGVDSSFFSYTGPS---DNNPWDEIDIEFLGKDT 147

Qy      67 GSFQSNIITGKAGAQKTSEKH---HAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
             | |  |     |      ||   : :   | : |||||  |  ||: | |||:|| :
Db     148 TEVQFNYYTNGVG------KHEYLYKLGFDASKGFHTYGYIWEQNYIAWLVDGKEVYRA- 200

Qy     124 GGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYXYTP 173
             ||:    | :  |||      :|:| :|    |:  : ||     | |
Db     201 ---TSNIPTHPGKVMMNLWPGIGVDSWLGAYD-GVTPVKAYYNWA---MYNP 245


RESULT 13
Q93UI1
ID   Q93UI1      PRELIMINARY;      PRT;   205 AA.
AC   Q93UI1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Endo 1-3,1-4-beta-glucanase (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rodriguez V., Mellado R.P.;
RT   "Novel lichenases from soil.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254960; AAK50613.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER     205    205
SQ   SEQUENCE   205 AA;  22890 MW;  E7CADAAFE02A2F51 CRC64;

  Query Match             16.4%;  Score 219;  DB 2;  Length 205;
  Best Local Similarity   36.2%;  Pred. No. 2.2e-10;
  Matches   55;  Conservative   20;  Mismatches   67;  Indels   10;  Gaps    5;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      55 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIESL 111

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
             ||:    | |  |  ||     ||   :   |  |:| |  :| || ::| |||| :: |
Db     112 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHMYAFDWQPNSIKWYVDGQ-LKHT 167

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 153
             ||    |   :  |||:       |:| ::
Db     168 ATTQVPAAPGK--IMMNLWNGTGVDDWLGSYN 197


RESULT 14
Q8U8N5
ID   Q8U8N5      PRELIMINARY;      PRT;   289 AA.
AC   Q8U8N5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Endo-1,3-1,4-beta-glycanase.
GN   EXOK OR ATU4055 OR AGR_L_1600.
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193;
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.;
RT   "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT   C58.";
RL   Science 294:2317-2323(2001).
RN   [2]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608551; PubMed=11743194;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
DR   EMBL; AE009336; AAL44856.1; ALT_INIT.
DR   EMBL; AE008279; AAK89373.1; -.
KW   Complete proteome.
SQ   SEQUENCE   289 AA;  32368 MW;  0384C6F83320EAC9 CRC64;

  Query Match             15.5%;  Score 206;  DB 16;  Length 289;
  Best Local Similarity   35.1%;  Pred. No. 4.1e-09;
  Matches   54;  Conservative   26;  Mismatches   56;  Indels   18;  Gaps    6;

Qy       5 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
           ::|:  |: |    :|| :||||| |  ||  |: | |   :   | :|  |:| |||||
Db     115 RNFACGEIQTKGRYRYGTYEARMKAATGSGLNSAFFTYIGPT---DKKPHDEIDFEVLGK 171

Qy      65 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTE 123
             | |  | |  | | |    :||   |   ||  |: |   | |  :|: |:|: |
Db     172 NTGKVQLNQYIAAKGG----NEKLVPVEGGADAGFNDYAFVWEPQRLRYYVNGKLVH--- 224

Qy     124 GGQVSNLT----GTQGLRFNLWSSESAA-WVGQF 152
             :|:: |      | : |:|| :::   |:|:|
Db     225 --EVTDETKIPQNAQKIFFSLWGTDTLKDWMGKF 256


RESULT 15
Q98C78
ID   Q98C78      PRELIMINARY;      PRT;   293 AA.
AC   Q98C78;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Endo-1,3-1,4-beta-glycanase, ExoK.
GN   MLR5264.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003006; BAB51743.1; -.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   PRINTS; PR00737; GLHYDRLASE16.
KW   Complete proteome.
SQ   SEQUENCE   293 AA;  32347 MW;  93BBA1672176ED3C CRC64;

  Query Match             15.0%;  Score 200;  DB 16;  Length 293;
  Best Local Similarity   33.8%;  Pred. No. 1.3e-08;
  Matches   51;  Conservative   27;  Mismatches   61;  Indels   12;  Gaps    6;

Qy       5 KDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
           ::|:  |: | :    || :|||:|    ||  :: | |   |   | :|| |:| |:| |
Db     112 REFACGEIQTKQRFGYGTYEARLKTDTGSGLNAAFFTYIGPS---DKQPWDEIDFEILTK 168

Qy      65 NPGSFQSN-IITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRK-T 122
           :    | |  | ||   :|  |    |    |:||: |   |  : :|| |:|| |   |
Db     169 DTSKVQVNAYIDGKGKNEKLVE----VPGGTDKAFNDYAFVWEKDSLRWYVNGQLVNTIT 224

Qy     123 EGGQVSNLTGTQGLRFNLWSSES-AAWVGQF 152
           :  ::   :  | : |:|| ||:   |:| |
Db     225 DPAKLP--SHAQKIFFSLWGSETMKGWMGAF 253


Search completed: January  9, 2003, 12:18:10
Job time : 24.1019 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:04:12 ; Search time 10.9074 Seconds
(without alignments)
1613.076 Million cell updates/sec

Title: US-09-654-652A-1
Perfect score: 1333
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....TRKGQESFNGQVPRDDEPAP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 263044 seqs, 70945482 residues

Total number of hits satisfying chosen parameters: 263044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	116.5	8.7	673	6	US-10-294-561-3	Sequence 3, Appli
2	104	7.8	1510	5	US-09-724-676-83340	Sequence 83340, A
3	104	7.8	1510	5	US-09-724-676A-83340	Sequence 83340, A
4	103.5	7.8	1554	5	US-09-724-676-83341	Sequence 83341, A
5	103.5	7.8	1554	5	US-09-724-676A-83341	Sequence 83341, A
6	99	7.4	1558	5	US-09-724-676-83343	Sequence 83343, A
7	99	7.4	1558	5	US-09-724-676A-83343	Sequence 83343, A
8	98.5	7.4	1602	5	US-09-724-676-83344	Sequence 83344, A
9	98.5	7.4	1602	5	US-09-724-676A-83344	Sequence 83344, A
10	97.5	7.3	1629	5	US-09-724-676-83342	Sequence 83342, A
11	97.5	7.3	1629	5	US-09-724-676A-83342	Sequence 83342, A
12	86.5	6.5	444	6	US-10-092-411A-4346	Sequence 4346, Ap
13	83.5	6.3	486	5	US-09-724-676-60984	Sequence 60984, A
14	83.5	6.3	486	5	US-09-724-676A-60984	Sequence 60984, A
15	83.5	6.3	638	5	US-09-917-384A-9	Sequence 9, Appli
16	83.5	6.3	640	5	US-09-917-384A-5	Sequence 5, Appli
17	83.5	6.3	1121	5	US-09-917-384A-1	Sequence 1, Appli
18	81	6.1	246	1	PCT-US02-36496-1920	Sequence 1920, Ap
19	81	6.1	246	6	US-10-293-418-1920	Sequence 1920, Ap
20	80.5	6.0	246	1	PCT-US02-36496-1264	Sequence 1264, Ap
21	80.5	6.0	246	6	US-10-293-418-1264	Sequence 1264, Ap
22	80.5	6.0	250	1	PCT-US02-36496-1280	Sequence 1280, Ap
23	80.5	6.0	250	6	US-10-293-418-1280	Sequence 1280, Ap
24	78.5	5.9	248	1	PCT-US02-36496-861	Sequence 861, App
25	78.5	5.9	248	1	PCT-US02-36496-1037	Sequence 1037, Ap
26	78.5	5.9	248	1	PCT-US02-36496-1248	Sequence 1248, Ap
27	78.5	5.9	248	1	PCT-US02-36496-1254	Sequence 1254, Ap
28	78.5	5.9	248	1	PCT-US02-36496-1267	Sequence 1267, Ap
29	78.5	5.9	248	6	US-10-293-418-861	Sequence 861, App
30	78.5	5.9	248	6	US-10-293-418-1037	Sequence 1037, Ap
31	78.5	5.9	248	6	US-10-293-418-1248	Sequence 1248, Ap
32	78.5	5.9	248	6	US-10-293-418-1254	Sequence 1254, Ap
33	78.5	5.9	248	6	US-10-293-418-1267	Sequence 1267, Ap
34	78	5.9	571	6	US-10-287-274-406	Sequence 406, App
35	78	5.9	1125	6	US-10-276-781-1219	Sequence 1219, Ap
36	77.5	5.8	247	1	PCT-US02-36496-1257	Sequence 1257, Ap
37	77.5	5.8	247	6	US-10-293-418-1257	Sequence 1257, Ap
38	77.5	5.8	248	1	PCT-US02-36496-1031	Sequence 1031, Ap
39	77.5	5.8	248	1	PCT-US02-36496-1040	Sequence 1040, Ap
40	77.5	5.8	248	1	PCT-US02-36496-1237	Sequence 1237, Ap
41	77.5	5.8	248	1	PCT-US02-36496-1246	Sequence 1246, Ap
42	77.5	5.8	248	1	PCT-US02-36496-1251	Sequence 1251, Ap
43	77.5	5.8	248	6	US-10-293-418-1031	Sequence 1031, Ap
44	77.5	5.8	248	6	US-10-293-418-1040	Sequence 1040, Ap
45	77.5	5.8	248	6	US-10-293-418-1237	Sequence 1237, Ap

ALIGNMENTS

RESULT 1
US-10-294-561-3
; Sequence 3, Application US/10294561
; GENERAL INFORMATION:
; APPLICANT: Seikagaku Corporation
; TITLE OF INVENTION: (1-3)-Beta-D-Glucan Binding Domain Protein, Measuring Method U
; TITLE OF INVENTION: Substance And Assay Kit
; FILE REFERENCE: Q72918
; CURRENT APPLICATION NUMBER: US/10/294,561
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: JP 2001-351943
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Tachypleus tridentatus
US-10-294-561-3

Query Match 8.7%; Score 116.5; DB 6; Length 673;
Best Local Similarity 22.8%; Pred. No. 0.014;
Matches 61; Conservative 43; Mismatches 91; Indels 73; Gaps 16;

```
Qy       1 MVSAK--DFSGAELYTLEEV--------QYGKFEARMKMAAASGTVSSMFLYQNGSEIAD 50
           :::||  |: | : ||   :         :||| ||:| : :  | |  ||
Db      73 VITAKREDYDGFK-YTSARLKTQFDKSWKYGKIEAKMAIPSFRG-VWVMF---------- 120

Qy      51 GRPWVEVDIEVLGKNPGSFQSNIITGK-AGAQKTSEKHHAVSPAADQA------------ 97
              |:  |     : | | : : |  :    :|     |  :|    |
Db     121 ---WMSGDNTNYVRWPSSGEIDFIEHRNTNNEKVRGTIHWSTPDGAHAHHNRESNTNGID 177

Qy      98 FHTYGLEWTPNYVRWTVDGQEV--RKTEGGQVSNLTGTQGLRFNLWSSESAA----WVGQ 151
           :| | :||  : |:| |:| :    | :||    : |    |  ::   : |    | |
Db     178 YHIYSVEWNSSIVKWFVNGNQYFEVKIQGG----VNGKSAFRNKVFVILNMAIGGNWPG- 232

Qy     152 FD--ESKLPLFQFINWVKVYKYTPGQGEGG-------------SDFTLDWTDNFDTFDG 195
           ||  :   |   :|::|:||:         |             |:  || ||  :  ||
Db     233 FDVADEAFPAKMYIDYVRVYQDASTSSPVGDTSLDGYYFVQNRHSELYLDVTDASNE-DG 291

Qy     196 SRWGKGDWTFDGNR-----VDLTDKNIY 218
           :   :  |:: ||       :  : |:|
Db     292 AFLQQ--WSYSGNENQQFDFEHLENNVY 317
```

RESULT 2
US-09-724-676-83340
; Sequence 83340, Application US/09724676
; GENERAL INFORMATION:

```
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83340
;   LENGTH: 1510
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83340

  Query Match             7.8%;  Score 104;  DB 5;  Length 1510;
  Best Local Similarity   19.7%;  Pred. No. 0.6;
  Matches   54;  Conservative   41;  Mismatches   95;  Indels   84;  Gaps   10;

Qy      12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     131 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 166
            |      |  :|: :: |                      ::| |  |: : :| :
Db    1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy     167 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR-------- 209
              :|||        | | | |:| :    |||:|    :|   |  | |
Db    1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARA 1423

Qy     210 ---------------VDLTDKNIYSRDGMLILA 227
                          :|:  :||  :  |   |
Db    1424 FVRKTSIFIMDEATASIDMATENILQKVVMTAFA 1457


RESULT 3
US-09-724-676A-83340
; Sequence 83340, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83340
;   LENGTH: 1510
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83340

  Query Match             7.8%;  Score 104;  DB 5;  Length 1510;
  Best Local Similarity   19.7%;  Pred. No. 0.6;
  Matches   54;  Conservative   41;  Mismatches   95;  Indels   84;  Gaps   10;

Qy      12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     131 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 166
            |      |  :|: :: |                      ::| |  |: : :| :
Db    1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy     167 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR-------- 209
              :|||        | | | |:| :    |||:|    :|   |  | |
Db    1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARA 1423

Qy     210 ---------------VDLTDKNIYSRDGMLILA 227
                          :|:  :||  :  |   |
Db    1424 FVRKTSIFIMDEATASIDMATENILQKVVMTAFA 1457


RESULT 4
US-09-724-676-83341
; Sequence 83341, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83341
;   LENGTH: 1554
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83341

  Query Match             7.8%;  Score 103.5;  DB 5;  Length 1554;
  Best Local Similarity   20.8%;  Pred. No. 0.7;
  Matches   46;  Conservative   36;  Mismatches   80;  Indels   59;  Gaps    8;

Qy      12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     131 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 166
            |      |  :|: :: |                      ::| |  |: : :| :
Db    1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy     167 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWG 199
              :|||        | | | |:| :    |||:|   |
Db    1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGQLEG 1404


RESULT 5
US-09-724-676A-83341
; Sequence 83341, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83341
;   LENGTH: 1554
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83341

  Query Match             7.8%;  Score 103.5;  DB 5;  Length 1554;
  Best Local Similarity   20.8%;  Pred. No. 0.7;
  Matches   46;  Conservative   36;  Mismatches   80;  Indels   59;  Gaps    8;

Qy      12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314
```

```
Qy    131 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 166
          |       |  :|: :: |                        ::| |  |: : :| :
Db   1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy    167 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWG 199
              :|||        | | | |:| :    |||:|   |
Db   1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGQLEG 1404


RESULT 6
US-09-724-676-83343
; Sequence 83343, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83343
;   LENGTH: 1558
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83343

  Query Match             7.4%;  Score 99;  DB 5;  Length 1558;
  Best Local Similarity   18.7%;  Pred. No. 1.8;
  Matches   59;  Conservative   38;  Mismatches  100;  Indels  118;  Gaps   11;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||            |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGT--------QGL-------RFNLWSSESAAWVG---------------------- 150
           |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                   ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    178 GGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR----------------------VDL 212
          | | |:| :    |||:|    :|   | | |                      :|:
Db   1431 GKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARAFVRKTSIFIMDEATASIDM 1490

Qy    213 TDKNIYSRDGMLILA 227
            :||  :  |   |
Db   1491 ATENILQKVVMTAFA 1505


RESULT 7
US-09-724-676A-83343
; Sequence 83343, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83343
;   LENGTH: 1558
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83343

  Query Match             7.4%;  Score 99;  DB 5;  Length 1558;
  Best Local Similarity   18.7%;  Pred. No. 1.8;
  Matches   59;  Conservative   38;  Mismatches  100;  Indels  118;  Gaps   11;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||            |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGT--------QGL-------RFNLWSSESAAWVG---------------------- 150
           |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                   ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    178 GGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR----------------------VDL 212
          | | |:| :    |||:|    :|   | | |                      :|:
Db   1431 GKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARAFVRKTSIFIMDEATASIDM 1490

Qy    213 TDKNIYSRDGMLILA 227
            :||  :  |   |
Db   1491 ATENILQKVVMTAFA 1505


RESULT 8
US-09-724-676-83344
; Sequence 83344, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83344
;   LENGTH: 1602
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83344

  Query Match             7.4%;  Score 98.5;  DB 5;  Length 1602;
  Best Local Similarity   19.5%;  Pred. No. 2.1;
  Matches   51;  Conservative   33;  Mismatches   85;  Indels   93;  Gaps    9;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||            |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGT--------QGL-------RFNLWSSESAAWVG---------------------- 150
           |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                   ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    178 GGSDFTLDWTDNFDTFDGSRWG 199
          | | |:| :    |||:|    |
Db   1431 GKSSFSLAFFRMVDTFEGQLEG 1452


RESULT 9
US-09-724-676A-83344
```

```
; Sequence 83344, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83344
;   LENGTH: 1602
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83344

  Query Match             7.4%;  Score 98.5;  DB 5;  Length 1602;
  Best Local Similarity   19.5%;  Pred. No. 2.1;
  Matches   51;  Conservative   33;  Mismatches   85;  Indels   93;  Gaps    9;

Qy      12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     131 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 150
            |         :||       |    | |   |||
Db    1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGGGTFGYSCGPCLVLSPA 1374

Qy     151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                    ::| |  |: : :| :     :|||        |
Db    1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy     178 GGSDFTLDWTDNFDTFDGSRWG 199
           | | |:| :    |||:|   |
Db    1431 GKSSFSLAFFRMVDTFEGQLEG 1452


RESULT 10
US-09-724-676-83342
; Sequence 83342, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83342
;   LENGTH: 1629
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83342

  Query Match             7.3%;  Score 97.5;  DB 5;  Length 1629;
  Best Local Similarity   19.4%;  Pred. No. 2.6;
  Matches   50;  Conservative   33;  Mismatches   82;  Indels   93;  Gaps    9;

Qy      12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     131 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 150
            |         :||       |    | |   |||
Db    1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGGGTFGYSCGPCLVLSPA 1374

Qy     151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                    ::| |  |: : :| :     :|||        |
Db    1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy     178 GGSDFTLDWTDNFDTFDG 195
           | | |:| :    |||:|
Db    1431 GKSSFSLAFFRMVDTFEG 1448


RESULT 11
US-09-724-676A-83342
; Sequence 83342, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83342
;   LENGTH: 1629
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83342

  Query Match             7.3%;  Score 97.5;  DB 5;  Length 1629;
  Best Local Similarity   19.4%;  Pred. No. 2.6;
  Matches   50;  Conservative   33;  Mismatches   82;  Indels   93;  Gaps    9;

Qy      12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     131 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 150
            |         :||       |    | |   |||
Db    1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGGGTFGYSCGPCLVLSPA 1374

Qy     151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                    ::| |  |: : :| :     :|||        |
Db    1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy     178 GGSDFTLDWTDNFDTFDG 195
           | | |:| :    |||:|
Db    1431 GKSSFSLAFFRMVDTFEG 1448


RESULT 12
US-10-092-411A-4346
; Sequence 4346, Application US/10092411A
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-101
;  CURRENT APPLICATION NUMBER: US/10/092,411A
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 09/134,001
;  PRIOR FILING DATE: 1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 4346
;   LENGTH: 444
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-10-092-411A-4346
```

```
  Query Match             6.5%;  Score 86.5;  DB 6;  Length 444;
  Best Local Similarity   19.8%;  Pred. No. 4.7;
  Matches   53;  Conservative   35;  Mismatches   71;  Indels  109;  Gaps   14;

Qy       2 VSAKDFSGA---ELYTLEEVQYGKF-EARMKMAAASGTV-SSMFLYQNGSEIADG----- 51
           :  |:| ||   ::   : : |  | : |  :    | :  :: :| || || ||
Db     232 MEGKNFKGAITIDMSESDSIAYESFKQKRFNVDTEGGYILDNVKIYANGKEIGDGFSSET 291

Qy      52 ----RPWVEVDIEVLGKNPG-SFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWT 106
               |  || :   |   | :|:|| :   : ::|
Db     292 YGPYDPDEEVIVHAEGSYEGKTFKSNSVNVASASEK---------------------- 327

Qy     107 PNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPLFQFINWV 166
                       :|| |:::|                  :|||  :   |:::
Db     328 ----------------DGG-VTDVT-----------------VKFDEEAID--QYVDKK 350

Qy     167 KVYKYTPGQGEGGSD--------------------FTLDWTDNFDTFDGSRWGKGDWTF- 205
             ||     | :|                     |||  || :   :  : | | | |
Db     351 LDEKYDDSDDESDNDSSSGEVTRENVIDKVESYEGHTLD-TDTYTYKEPEKTGDGKWGFS 409

Qy     206 ----DGN-----RVDLTDKNI--YSRDG 222
               ||:      ||: |  :  |  ||
Db     410 FLDKDGDLAGSYTVDIDDGYVTEYDEDG 437


RESULT 13
US-09-724-676-60984
; Sequence 60984, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60984
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-60984

  Query Match             6.3%;  Score 83.5;  DB 5;  Length 486;
  Best Local Similarity   22.1%;  Pred. No. 9.9;
  Matches   30;  Conservative   21;  Mismatches   48;  Indels   37;  Gaps    4;

Qy     103 LEWTPNYVRWTVDGQEVRKTEGGQVSNL-----------------TGTQGLRFNLWSSE 144
           | |   |::  :  :::   |  : |:|                 ||:  |   ||
Db     149 LHWKKVYLKAILRMKQLEDHEAFETSSLIGHSARVYALYYKDGLLCTGSDDLSAKLWDVS 208

Qy     145 SAAWV----------GQFDESKLPLFQFINWVKVYKYTPG------QGEGGSDFTLDWTD 188
           :   |           :||| ||    | | |  :::: |      :|  |: |::|: |
Db     209 TGQCVYGIQTHTCAAVKFDEQKLVTGSFDNTVACWEWSSGARTQHFRGHTGAVFSVDYND 268

Qy     189 NFDTFDGSRWGKGDWT 204
             |       |  |:|
Db     269 ELDILVS---GSADFT 281


RESULT 14
US-09-724-676A-60984
; Sequence 60984, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60984
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-60984

  Query Match             6.3%;  Score 83.5;  DB 5;  Length 486;
  Best Local Similarity   22.1%;  Pred. No. 9.9;
  Matches   30;  Conservative   21;  Mismatches   48;  Indels   37;  Gaps    4;

Qy     103 LEWTPNYVRWTVDGQEVRKTEGGQVSNL-----------------TGTQGLRFNLWSSE 144
           | |   |::  :  :::   |  : |:|                 ||:  |   ||
Db     149 LHWKKVYLKAILRMKQLEDHEAFETSSLIGHSARVYALYYKDGLLCTGSDDLSAKLWDVS 208

Qy     145 SAAWV----------GQFDESKLPLFQFINWVKVYKYTPG------QGEGGSDFTLDWTD 188
           :   |           :||| ||    | | |  :::: |      :|  |: |::|: |
Db     209 TGQCVYGIQTHTCAAVKFDEQKLVTGSFDNTVACWEWSSGARTQHFRGHTGAVFSVDYND 268

Qy     189 NFDTFDGSRWGKGDWT 204
             |       |  |:|
Db     269 ELDILVS---GSADFT 281


RESULT 15
US-09-917-384A-9
; Sequence 9, Application US/09917384A
; GENERAL INFORMATION:
;  APPLICANT: ADNEY, WILLIAM S.
;  APPLICANT:  DING, SHI-YOU
;  APPLICANT:  MCCARTER, SUZANNE
;  APPLICANT:  HIMMEL, MICHAEL E.
;  APPLICANT:  DECKER, STEPHEN R.
;  APPLICANT:  VINZANT, TODD B.
;  TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
;  TITLE OF INVENTION:  CELLULOLYTICUS
;  FILE REFERENCE: NREL 01-38
;  CURRENT APPLICATION NUMBER: US/09/917,384A
;  CURRENT FILING DATE:  2001-07-28
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 638
;   TYPE: PRT
;   ORGANISM: Acidothermus cellulolyticus
US-09-917-384A-9

  Query Match             6.3%;  Score 83.5;  DB 5;  Length 638;
  Best Local Similarity   22.2%;  Pred. No. 14;
  Matches   36;  Conservative   21;  Mismatches   56;  Indels   49;  Gaps    7;

Qy      79 GAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRF 138
           ||  :   ::   || |  |:    :| | |                   :|  |     |
Db     375 GATNSWNGNYGTPPAGDSTFYGMAYDWEPVY-------------HDPPSNNWFG-----F 416

Qy     139 NLWSSESAA---WVGQFDESKLPLFQFINWVKVYKYTPGQGEGGS---DFTLDWTDNFDT 192
             || |  |   :|    ::|  | ::: |||     |    | |     | |:   ||
Db     417 QAWSMERVAEYYYVTGDPKAKALLDKWVAWVK-----PNVTTGASWSIPSNLSWSGQPDT 471

Qy     193 FDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQE 234
           :: |  |             |: |::       : :|  ||:
Db     472 WNPSNPG-------------TNANLH-------VTITSSGQD 493


Search completed: January  9, 2003, 12:14:11
Job time : 12.9074 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:16 ; Search time 6.31481 Seconds
(without alignments)
761.932 Million cell updates/sec

Title: US-09-654-652A-1
Perfect score: 1333
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....TRKGQESFNGQVPRDDEPAP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	120	9.0	287	10	US-09-734-569-152	Sequence 152, App
2	117	8.8	620	10	US-09-988-200-6	Sequence 6, Appli
3	113	8.5	545	10	US-09-988-200-8	Sequence 8, Appli
4	86.5	6.5	1385	10	US-09-738-363-2	Sequence 2, Appli
5	84.5	6.3	1289	10	US-09-738-363-4	Sequence 4, Appli
6	84	6.3	714	9	US-09-738-626-6377	Sequence 6377, Ap
7	83.5	6.3	422	12	US-10-042-417-4	Sequence 4, Appli
8	83.5	6.3	848	10	US-09-841-132-192	Sequence 192, App
9	83.5	6.3	1530	10	US-09-841-132-178	Sequence 178, App
10	83.5	6.3	1531	12	US-10-007-693-98	Sequence 98, Appl
11	80.5	6.0	467	10	US-09-863-547B-1	Sequence 1, Appli
12	80	6.0	156	10	US-09-734-569-28	Sequence 28, Appl
13	78.5	5.9	475	9	US-09-738-626-6052	Sequence 6052, Ap
14	77.5	5.8	231	10	US-09-728-911-2	Sequence 2, Appli
15	77.5	5.8	231	10	US-09-949-192-6	Sequence 6, Appli
16	77	5.8	200	10	US-09-734-569-150	Sequence 150, App
17	77	5.8	1233	9	US-09-738-626-4312	Sequence 4312, Ap
18	75.5	5.7	165	10	US-09-925-299-865	Sequence 865, App
19	75.5	5.7	731	9	US-09-738-626-4854	Sequence 4854, Ap
20	75.5	5.7	1375	10	US-09-740-274-4	Sequence 4, Appli
21	75	5.6	245	9	US-09-738-626-4876	Sequence 4876, Ap
22	75	5.6	2167	10	US-09-778-927A-61	Sequence 61, Appl
23	75	5.6	2201	12	US-10-100-912-2	Sequence 2, Appli
24	74	5.6	925	10	US-09-452-380-4	Sequence 4, Appli
25	74	5.6	936	10	US-09-452-380-3	Sequence 3, Appli
26	73.5	5.5	356	10	US-09-289-346A-11	Sequence 11, Appl
27	73.5	5.5	678	10	US-09-801-368-314	Sequence 314, App
28	73.5	5.5	862	10	US-09-833-435A-1	Sequence 1, Appli
29	73.5	5.5	930	10	US-09-815-242-13481	Sequence 13481, A
30	72.5	5.4	707	10	US-09-764-870-313	Sequence 313, App
31	72	5.4	395	9	US-10-094-080-3	Sequence 3, Appli
32	72	5.4	433	9	US-09-964-899-11	Sequence 11, Appl
33	72	5.4	511	9	US-09-738-626-5606	Sequence 5606, Ap
34	72	5.4	700	10	US-09-854-122-50	Sequence 50, Appl
35	72	5.4	1054	9	US-09-738-626-5852	Sequence 5852, Ap
36	71.5	5.4	406	10	US-09-215-450-26	Sequence 26, Appl
37	71.5	5.4	663	10	US-09-815-242-5350	Sequence 5350, Ap
38	71.5	5.4	663	10	US-09-815-242-12138	Sequence 12138, A
39	71.5	5.4	915	10	US-09-332-226-2	Sequence 2, Appli
40	71.5	5.4	1050	10	US-09-866-562-38	Sequence 38, Appl
41	71	5.3	345	10	US-09-925-300-1188	Sequence 1188, Ap
42	71	5.3	350	9	US-09-738-626-5115	Sequence 5115, Ap
43	71	5.3	420	10	US-09-905-176-25	Sequence 25, Appl
44	71	5.3	471	10	US-09-765-272-158	Sequence 158, App
45	71	5.3	486	10	US-09-815-242-13455	Sequence 13455, A

ALIGNMENTS

RESULT 1
US-09-734-569-152
; Sequence 152, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv
; TITLE OF INVENTION: in the synthesis of carbohydrates
; FILE REFERENCE: BASF-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734,569
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 152
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-152

Query Match 9.0%; Score 120; DB 10; Length 287;
Best Local Similarity 23.0%; Pred. No. 0.00037;
Matches 49; Conservative 30; Mismatches 70; Indels 64; Gaps 11;

```
Qy      20 YGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGS---FQSNII 74
           |    | :||    ::|||:: ::   | :        |:|:| ||   |      :|:
Db      69 YVDISAYIKMPPFDSAGTVTTFYMSSQGDQ------HYELDMEFLGNTSGQPFLLHTNVF 122

Qy      75 TGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGT- 133
              | ::  : :    |:|| || |   |: : | : || : ||          || ||
Db     123 VDGVGGRE-QQMYLGFDPSAD--FHYYRFRWSKDMVVFYVDNKPVR-----VFKNLEGTV 174
```

```
Qy    134 ----------QGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFT 183
                  |:  ::|   | |  |     ::|    |||             : ||
Db    175 PGTKYLNQQAMGVYISIWDGSSWATQG----GRVP----INW-----------ASAPFT 214

Qy    184 LDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKN 216
            :              |:  :|  ||  | |
Db    215 ATYQ-------------DFALNGCVVDPNDPN 233


RESULT 2
US-09-988-200-6
; Sequence 6, Application US/09988200
; Patent No. US20020094553A1
;   GENERAL INFORMATION:
;        APPLICANT: BARBEYRON, Tristan
;                   POTIN, Philippe
;                   RICHARD, Christophe
;                   HENRISSAT, Bernard
;                   YVIN, Jean-Claude
;                   KLOAREG, Bernard
;        TITLE OF INVENTION: Glycolyse hydrolase genes and their
;                            use for producing enzymes for the biodegradation of
;                            carrageenans
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DENNISON, MESEROLE, SCHEINER & SCHULTZ
;             STREET: 612 Crystal Square 4, 1745 Jefferson Davis
;                     Highway
;             CITY: ARLINGTON
;             STATE: VIRGINIA
;             COUNTRY: U.S.A.
;             ZIP: 22202
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/988,200
;             FILING DATE: 19-No. US20020094553A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/269,731
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: FR 96 12204
;             FILING DATE: 07-OCT-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: IRA SCHULTZ
;             REGISTRATION NUMBER: <Unknown>
;             REFERENCE/DOCKET NUMBER: <Unknown>
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 412-1155
;             TELEFAX: (703) 412-1161
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 620 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-988-200-6

  Query Match             8.8%;  Score 117;  DB 10;  Length 620;
  Best Local Similarity   27.3%;  Pred. No. 0.002;
  Matches   54;  Conservative   23;  Mismatches   73;  Indels   48;  Gaps   10;

Qy      20 YGKFEARMKMAAASGTVSSMF-LYQ--NGSEIADGR-PWVEVDI---------------- 59
           ||.:|||:| |:    ||  | :|   : |   :|   : |:|:
Db     289 YGYYEARIKGASTFPGVSPAFWMYSTIDRSLTKEGDVQYSEIDVVELTQKSAVRESDHDL 348

Qy      60 -EVLGKN-------PGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVR 111
            :: ||       ||||      |          |    |  |  |||||:  | : :
```

```
Db    349 HNIVVKNGKPTWMRPGSFPQTNHNG---------YHLPFDPRND--FHTYGVNVTKDKIT 397

Qy    112 WTVDGQEVRKTEG---GQVSNLTGTQGLR--FNLWSSE----SAAWVGQFDESKLPLFQF 162
          | |||: | : :      :  ||| :|||     |       ||    :   : : :
Db    398 WYVDGEIVGEKDNLYWHRQMNLTLSQGLRAPHTQWKCNQFYPSANKSAEGFPTSMEVDYV 457

Qy    163 INWVKVYKYTPGQGEGGS 180
            ||||       ||| |
Db    458 RTWVKVGNNNSAPGEGQS 475


RESULT 3
US-09-988-200-8
; Sequence 8, Application US/09988200
; Patent No. US20020094553A1
;   GENERAL INFORMATION:
;        APPLICANT: BARBEYRON, Tristan
;                   POTIN, Philippe
;                   RICHARD, Christophe
;                   HENRISSAT, Bernard
;                   YVIN, Jean-Claude
;                   KLOAREG, Bernard
;        TITLE OF INVENTION: Glycolyse hydrolase genes and their
;                            use for producing enzymes for the biodegradation of
;                            carrageenans
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DENNISON, MESEROLE, SCHEINER & SCHULTZ
;             STREET: 612 Crystal Square 4, 1745 Jefferson Davis
;                     Highway
;             CITY: ARLINGTON
;             STATE: VIRGINIA
;             COUNTRY: U.S.A.
;             ZIP: 22202
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/988,200
;             FILING DATE: 19-No. US20020094553A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/269,731
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: FR 96 12204
;             FILING DATE: 07-OCT-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: IRA SCHULTZ
;             REGISTRATION NUMBER: <Unknown>
;             REFERENCE/DOCKET NUMBER: <Unknown>
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 412-1155
;             TELEFAX: (703) 412-1161
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 545 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-988-200-8

  Query Match             8.5%;  Score 113;  DB 10;  Length 545;
  Best Local Similarity   23.0%;  Pred. No. 0.004;
  Matches   47;  Conservative   41;  Mismatches   78;  Indels   38;  Gaps   10;

Qy       7 FSGAELYTLEEVQYGKFEARMKMA-AASGTVSSMFLYQN-GSEIADGRP-WVEVDIEVL- 62
           |:     : ::  || |||::: |    |   | :|| :    :|:|   : |:|:  |
Db     108 FTSGIFKSYQKFTYGYFEAKIQGADIGEGVCPSFWLYSDFDYSVANGETVYSEIDVVELQ 167

Qy      63 ------GKNPGSFQSNI----ITGKAG----------AQKTSEKHHAVSPAADQAFHTYG 102
```

```
Db    168 QFDWYEGHQDDIYDMDLNLHAVVKENGQGVWKRPKMYPQEQLNKWRAMDPSKD--FHIYG 225

Qy    103 LEWTPNYVRWTVDGQEVRKTEG---GQVSNLTGTQGLR---FNLWSSESAAWVGQFDE-- 154
Db    226 CEVNQNEIIWYVDGVEVARKPNKYWHRPMNVTLSLGLRKPFVKFFDNKNNAINPETDAKA 285

Qy    155 ----SKLPLFQFINWVKVYKYTPG 174
Db    286 REKLSDIPTSMYVDYVRVWEKSAG 309


RESULT 4
US-09-738-363-2
; Sequence 2, Application US/09738363
; Patent No. US20010010932A1
;   GENERAL INFORMATION:
;        APPLICANT: Schnepf, Harry E.
;                   Schwab, George E.
;                   Payne, Jewel M.
;                   Narva, Kenneth E.
;                   Foncerrada, Luis
;        TITLE OF INVENTION: Nematicidal Proteins
;        NUMBER OF SEQUENCES: 40
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Jay M. Sanders
;             STREET: 2421 N.W. 41st Street, Suite A-1
;             CITY: Gainesville
;             STATE: FL
;             COUNTRY: USA
;             ZIP: 32606
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/738,363
;             FILING DATE: 15-Dec-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/076,137
;             FILING DATE: 12-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sanders, Jay
;             REGISTRATION NUMBER: 39,355
;             REFERENCE/DOCKET NUMBER: MA-20CCCD3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 352-375-8100
;             TELEFAX: 352-372-5800
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1385 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             ORGANISM: BACILLUS THURINGIENSIS
;             INDIVIDUAL ISOLATE: PS17
;        IMMEDIATE SOURCE:
;             CLONE: E. coli NM522(pMYC 1627) NRRL B-18651
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-738-363-2

  Query Match             6.5%;  Score 86.5;  DB 10;  Length 1385;
  Best Local Similarity   24.0%;  Pred. No. 4.8;
  Matches   56;  Conservative   29;  Mismatches   85;  Indels   63;  Gaps   13;

Qy     33 SGTVSSMFLYQNGSEIA-----DGRPWVEVD-IEVLGKNPGSFQSNIITGKAGAQKTSEK 86
Db   1176 NGSIRSDISYQNIDAIVLPTLPKLRHWFMSDRFSEQGDIMAKFQGALNRAYAQLEQSTLL 1235

Qy     87 HHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESA 146
Db   1236 HNG-------------HFTKDAANWTIEG------DAHQITLEDGRRVLRLPDWSSS-- 1273

Qy    147 AWVGQFDESKLPLFQFINWVKVYKYT-PGQGEGGSDFTLD----------WTDNFDTFDG 195
Db   1274 --VSQMIE-----IENFNPDKEYNLVFHGQGEG--TVTLEHGEETKYIETHTHHFANFTT 1324

Qy    196 SRWGKGDWTFDGNRVDLTDKNIYSRDGMLI---LALTRKGQESFNGQVPRDDE 245
Db   1325 SQ--RQGLTFESNKVTVT---ISSEDGEFLVDNIALV-------EAPLPTDDQ 1365


RESULT 5
US-09-738-363-4
; Sequence 4, Application US/09738363
; Patent No. US20010010932A1
;   GENERAL INFORMATION:
;        APPLICANT: Schnepf, Harry E.
;                   Schwab, George E.
;                   Payne, Jewel M.
;                   Narva, Kenneth E.
;                   Foncerrada, Luis
;        TITLE OF INVENTION: Nematicidal Proteins
;        NUMBER OF SEQUENCES: 40
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Jay M. Sanders
;             STREET: 2421 N.W. 41st Street, Suite A-1
;             CITY: Gainesville
;             STATE: FL
;             COUNTRY: USA
;             ZIP: 32606
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/738,363
;             FILING DATE: 15-Dec-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/076,137
;             FILING DATE: 12-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sanders, Jay
;             REGISTRATION NUMBER: 39,355
;             REFERENCE/DOCKET NUMBER: MA-20CCCD3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 352-375-8100
;             TELEFAX: 352-372-5800
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1289 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             ORGANISM: BACILLUS THURINGIENSIS
;             INDIVIDUAL ISOLATE: PS17
;        IMMEDIATE SOURCE:
;             CLONE: E. coli NM522(pMYC 1628) NRRL B-18652
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-738-363-4

  Query Match             6.3%;  Score 84.5;  DB 10;  Length 1289;
  Best Local Similarity   23.9%;  Pred. No. 6.8;
  Matches   56;  Conservative   27;  Mismatches   86;  Indels   65;  Gaps   13;
```

```
Qy      33 SGTVSSMFLYQNGSEIA-----DGRPWVEVD-IEVLGKNPGSFQSNIITGKAGAQKTSEK 86
           :|:: |   |||   |        | |   |      |     ||  :    |  :: :
Db    1080 NGSIRSDISYQNIDAIVLPTLPKLRHWFMSDRFSEQGDIMAKFQGALNRAYAQLEQNTLL 1139

Qy      87 HHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSS--E 144
           |:                :| :   |||:|      :  ||    | : ||   |||
Db    1140 HNG--------------HFTKDAANWTVEG------DAHQVVLEDGKRVLRLPDWSSSVS 1179

Qy     145 SAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLD----------WTDNFDTFD 194
               :  ||  |   :| :          |||||    ||:           | :|  |
Db    1180 QTIEIENFDPDK--EYQLV--------FHGQGEG--TVTLEHGEETKYIETHTHHFANFT 1227

Qy     195 GSRWGKGDWTFDGNRVDLTDKNIYSRDGMLI---LALTRKGQESFNGQVPRDDE 245
            |:  :   ||: |:| :|   | | ||  :   :||          :| ||:
Db    1228 TSQ--RQGLTFESNKVTVT---ISSEDGEFLVDNIALV-------EAPLPTDDQ 1269


RESULT 6
US-09-738-626-6377
; Sequence 6377, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6377
;   LENGTH: 714
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-6377

  Query Match             6.3%;  Score 84;  DB 9;  Length 714;
  Best Local Similarity   22.5%;  Pred. No. 3.5;
  Matches   51;  Conservative   36;  Mismatches  102;  Indels   38;  Gaps   11;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAAS-----GTVSSMFLYQNGSEIA--DGRPWVEVDI 59
           :| | :: | |:  | |   ::  | |     |   ::::  | :  |  ||||: |:||
Db     436 YSSARVH-LPEIPAGNFRLTVRARAQSEELVDGVRPAIWMQNNTNFCADNDGRPYGELDI 494

Qy      60 -EVLGKNPGSFQSNIITGKAG---AQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVD 115
            |       : | :  | ||     |  :     |   |  :| :|:|     : :|:|
Db     495 TEFYSSRVNTQYSAVHLGCAGNRPEMKLRQMEMEESMFGD--WHDWGVEVFDGQIVFTID 552

Qy     116 GQEV----RKTEGGQVSNLTG-TQGLRFNLWSSESAAWVGQFDESKLPLFQFINWVKVYK 170
           |: |    :   |  |:      :   | |   |    :||      |    :| :
Db     553 GKAVTSSGKDVFGNSVTPAAAPLRPAHFKLSEEEYREVIGQ------PWHLILNTM---- 602

Qy     171 YTPGQGEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNI 217
               : | |  :   || :-|   |      :  |   ||:   ::
Db     603 ----VEQSGKDSWITAVDNNEAFPEHR-----FQIDHVAVDIESDSV 640


RESULT 7
```

```
US-10-042-417-4
; Sequence 4, Application US/10042417
; Patent No. US20020123082A1
; GENERAL INFORMATION:
;  APPLICANT: Pagano, M.
;  TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
;  TITLE OF INVENTION:  PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
;  FILE REFERENCE: 5914-090-999
;  CURRENT APPLICATION NUMBER: US/10/042,417
;  CURRENT FILING DATE:  2002-01-07
;  PRIOR APPLICATION NUMBER: 60/260,179
;  PRIOR FILING DATE: 2001-01-5
;  NUMBER OF SEQ ID NOS: 89
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 422
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-042-417-4

  Query Match             6.3%;  Score 83.5;  DB 12;  Length 422;
  Best Local Similarity   22.1%;  Pred. No. 2;
  Matches   30;  Conservative   21;  Mismatches   48;  Indels   37;  Gaps    4;

Qy     103 LEWTPNYVRWTVDGQEVRKTEGGQVSNL-----------------TGTQGLRFNLWSSE 144
           | |   |::  :  :::   |  : |:|                  ||:  |   ||
Db     117 LHWKKVYLKAILRMKQLEDHEAFETSSLIGHSARVYALYYKDGLLCTGSDDLSAKLWDVS 176

Qy     145 SAAWV----------GQFDESKLPLFQFINWVKVYKYTPG------QGEGGSDFTLDWTD 188
           :   |           :||| ||    | | |  :::: |      :|  |: |::|: |
Db     177 TGQCVYGIQTHTCAAVKFDEQKLVTGSFDNTVACWEWSSGARTQHFRGHTGAVFSVDYND 236

Qy     189 NFDTFDGSRWGKGDWT 204
             |       |  |:|
Db     237 ELDILVS---GSADFT 249


RESULT 8
US-09-841-132-192
; Sequence 192, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 192
;   LENGTH: 848
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-841-132-192

  Query Match             6.3%;  Score 83.5;  DB 10;  Length 848;
  Best Local Similarity   20.4%;  Pred. No. 4.9;
  Matches   50;  Conservative   34;  Mismatches   82;  Indels   79;  Gaps   12;

Qy       8 SGAELYTLEEVQYG-------------KFEARMKMAAASGTVSSMFL-----YQNGSEIA 49
           :| ::  :|:   |              ||:| :      |:| : ||     ::    :
Db     599 AGVDIQLMEDFVLGVSGAAFLGKMDSQKFDAEVSRKGVVGSVYTGFLAGSWFFKGQYSLG 658

Qy      50 DGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNY 109
           : :  ::    |||::  |: |  :      |    |    | |        | | : | |
Db     659 ETQNDMKTRYGVLGESSASWTSRGVL----ADALVEYRSLVGPVRPT---FYALHFNP-Y 710

Qy     110 VR----------WTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPL 159
           |           :|  |:| |  |   ::|:|   |::| |      |   ||| |
Db     711 VEVSYASMKFPGFTEQGREARSFEDASLTNITIPLGMKFEL-----AFIKGQFSE----- 760
```

```
Qy    160 FQFIN-------WVKVYKYTPGQG----EGGSDF-----------------TLDWTDNF 190
          :|       | : |:   |      | | |:                    :|:  |
Db    761 ---VNSLGISYAW-EAYRKVEGGAVQLLEAGFDWEGAPMDLPRQELRVALENNTEWSSYF 816

Qy    191 DTFDG 195
          |  |
Db    817 STVLG 821
```

RESULT 9
US-09-841-132-178
; Sequence 178, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 178
; LENGTH: 1530
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-841-132-178

Query Match 6.3%; Score 83.5; DB 10; Length 1530;
Best Local Similarity 20.4%; Pred. No. 11;
Matches 50; Conservative 34; Mismatches 82; Indels 79; Gaps 12;

```
Qy      8 SGAELYTLEEVQYG-------------KFEARMKMAAASGTVSSMFL-----YQNGSEIA 49
          :| ::  :|:  |              ||:| :     |:| : ||     ::    :
Db   1281 AGVDIQLMEDFVLGVSGAAFLGKMDSQKFDAEVSRKGVVGSVYTGFLAGSWFFKGQYSLG 1340

Qy     50 DGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNY 109
          : :  ::    |||::  |: |  :     |    |   . | |         | | : | |
Db   1341 ETQNDMKTRYGVLGESSASWTSRGVL----ADALVEYRSLVGPVRPT---FYALHFNP-Y 1392

Qy    110 VR----------WTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPL 159
          |           :|  |:| |  |   ::|:|   |::| |     |   ||| |
Db   1393 VEVSYASMKFPGFTEQGREARSFEDASLTNITIPLGMKFEL-----AFIKGQFSE----- 1442

Qy    160 FQFIN-------WVKVYKYTPGQG----EGGSDF-----------------TLDWTDNF 190
          :|       | : |:   |      | | |:                    :|:  |
Db   1443 ---VNSLGISYAW-EAYRKVEGGAVQLLEAGFDWEGAPMDLPRQELRVALENNTEWSSYF 1498

Qy    191 DTFDG 195
          |  |
Db   1499 STVLG 1503
```

RESULT 10
US-10-007-693-98
; Sequence 98, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 98
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D

US-10-007-693-98

Query Match 6.3%; Score 83.5; DB 12; Length 1531;
Best Local Similarity 20.4%; Pred. No. 11;
Matches 50; Conservative 34; Mismatches 82; Indels 79; Gaps 12;

```
Qy      8 SGAELYTLEEVQYG-------------KFEARMKMAAASGTVSSMFL-----YQNGSEIA 49
          :| ::  :|:  |              ||:| :     |:| : ||     ::    :
Db   1282 AGVDIQLMEDFVLGVSGAAFLGKMDSQKFDAEVSRKGVVGSVYTGFLAGSWFFKGQYSLG 1341

Qy     50 DGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNY 109
          : :  ::    |||::  |: |  :     |    |   | |         | | : | |
Db   1342 ETQNDMKTRYGVLGESSASWTSRGVL----ADALVEYRSLVGPVRPT---FYALHFNP-Y 1393

Qy    110 VR----------WTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFDESKLPL 159
          |           :|  |:| |  |   ::|:|   |::| |     |   ||| |
Db   1394 VEVSYASMKFPGFTEQGREARSFEDASLTNITIPLGMKFEL-----AFIKGQFSE----- 1443

Qy    160 FQFIN-------WVKVYKYTPGQG----EGGSDF-----------------TLDWTDNF 190
          :|       | : |:   |      | | |:                    :|:  |
Db   1444 ---VNSLGISYAW-EAYRKVEGGAVQLLEAGFDWEGAPMDLPRQELRVALENNTEWSSYF 1499

Qy    191 DTFDG 195
          |  |
Db   1500 STVLG 1504
```

RESULT 11
US-09-863-547B-1
; Sequence 1, Application US/09863547B
; Patent No. US20020128166A1
; GENERAL INFORMATION:
; APPLICANT: Henkel KGaA
; APPLICANT: Hermanus, Lentus B. M.
; APPLICANT: Van Beckhoven, Rudolf F. W. C.
; APPLICANT: Maurer, Karl-Heinz
; APPLICANT: Kottwitz, Beatrix
; APPLICANT: Weiss, Albrecht
; APPLICANT: Van Solingen, Pieter
; TITLE OF INVENTION: Detergents Comprising Cellulases
; FILE REFERENCE: H 1920 A
; CURRENT APPLICATION NUMBER: US/09/863,547B
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 08/945,574
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: PCT/EP96/01755
; PRIOR FILING DATE: 1996-04-26
; PRIOR APPLICATION NUMBER: US 08/614,115
; PRIOR FILING DATE: 1996-03-12
; PRIOR APPLICATION NUMBER: EP 95201115.3
; PRIOR FILING DATE: 1995-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Bacillus sp. CBS 670.93
US-09-863-547B-1

Query Match 6.0%; Score 80.5; DB 10; Length 467;
Best Local Similarity 25.8%; Pred. No. 4.4;
Matches 42; Conservative 20; Mismatches 66; Indels 35; Gaps 11;

```
Qy     44 NGSEIA---DGRPWVEVDIEVLGKNPGSFQSNIITGKAGAQKTSEKHHA-----VSPAAD 95
          |||::      :|: |  | |:  |     :||:    |   : : |||       |
Db    167 NGSDVTWDNQIKPYAEEVIPVIRDND---PNNIVIVGTGTW-SQDVHHAADNQLADPNVM 222

Qy     96 QAFHTYGLEWTPN---YVRWTVD-GQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQ 151
          ||| |      |    | : :| |  :  :| | |  |  ||  |    ::  |:  |:
Db    223 YAFHFYAGTHGQNLRDQVDYALDQGAAIFVSEWG-TSAATGDGG----VFLDEAQVWIDF 277

Qy    152 FDESKLPLFQFINWVKVYK------YTPGQGEGGSDFTLDWTD 188
          ||  |    : ||   :|        ||    |      ||:
```

```
Db    278 MDERNL---SWANWSLTHKDESSAALMPGANPTGG-----WTE 312


RESULT 12
US-09-734-569-28
; Sequence 28, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
;  APPLICANT: Lerchl, Jens
;  APPLICANT:  Renz, Andreas
;  APPLICANT:  Ehrhardt, Thomas
;  APPLICANT:  Reindl, Andreas
;  APPLICANT:  Cirpus, Petra
;  APPLICANT:  Bischoff, Friedrich
;  APPLICANT:  Frank, Markus
;  APPLICANT:  Freund, Annette
;  APPLICANT:  Duwenig, Elke
;  APPLICANT:  Schmidt, Ralf-Michael
;  APPLICANT:  Reski, Ralf
;  TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
;  TITLE OF INVENTION:  in the synthesis of carbohydrates
;  FILE REFERENCE: BASF-NAE-1332-99-US
;  CURRENT APPLICATION NUMBER: US/09/734,569
;  CURRENT FILING DATE:  2001-05-24
;  PRIOR APPLICATION NUMBER: US 60/171,101
;  PRIOR FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 181
;  SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 28
;   LENGTH: 156
;   TYPE: PRT
;   ORGANISM: Physcomitrella patens
US-09-734-569-28

  Query Match             6.0%;  Score 80;  DB 10;  Length 156;
  Best Local Similarity   25.0%;  Pred. No. 1.2;
  Matches   24;  Conservative   19;  Mismatches   39;  Indels   14;  Gaps    5;

Qy      20 YGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGS---FQSNII 74
           |    | :||    ::|||:: ::   | :         |:|:| ||   |      :|:
Db      69 YVDISAYIKMPPFDSAGTVTTFYMSSQGDQ------HYELDMEFLGNTSGQPFLLHTNVF 122

Qy      75 TGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYV 110
              | ::  : :    |:||  || |   |: : |
Db     123 VDGVGGRE-QQMYLGFDPSAD--FHYYRFRWSKDMV 155


RESULT 13
US-09-738-626-6052
; Sequence 6052, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6052
;   LENGTH: 475
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-6052

  Query Match             5.9%;  Score 78.5;  DB 9;  Length 475;
  Best Local Similarity   24.5%;  Pred. No. 7;
  Matches   59;  Conservative   25;  Mismatches   94;  Indels   63;  Gaps   14;

Qy      14 TLEEVQYGKFEARMKMAAASGTVSSMFLYQNG---SEIADGRPWVEVDIEVLG------- 63
           | :  |||   |       |     ::| : |   ||:|    |: | : | |
Db     142 TGDTTQYGSSAALPFRQWVSQGHRMLWLGEPGRIYSELA--ASWLGV-LAVGGFALLWLR 198

Qy      64 -KNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQ--AFHTY-GLEWTP----------NY 109
            | ||  :  : ||  |  ||  :| |:   |     | |: || |:
Db     199 NKKPGRLRKMVRTGGRGRVKTYRRHAALGTVAGLGFVFLTFTGLTWSTYAGSNITDLRTQ 258

Qy     110 VRWTVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSESAAWVGQFD--------ESKLPLFQ 161
           : ||        |   ||       |   :: |: | :  |  |        | : ||
Db     259 LNWTQPSVNASLTAAPQVDMHDEHAGHHMHMESATSGS--GSIDLVAATAISELRTPL-- 314

Qy     162 FINWVKVYKYTPGQGEGGSDFTLDW--TDNFDTFDGSRWGKGDWTFDG------NRVDLT 213
                 ||   :|     | |  |:| | :   |:       ||      ||:: |
Db     315 --------TITPPAQDG-----LAWTATENRDAY---RFTTDTIAVDGDTGMLTNRLNST 358

Qy     214 D 214
           |
Db     359 D 359


RESULT 14
US-09-728-911-2
; Sequence 2, Application US/09728911
; Patent No. US20020012669A1
; GENERAL INFORMATION:
;  APPLICANT:  Presnell, Scott R.
;  APPLICANT:  Xu, Wenfeng
;  APPLICANT:  Kindsvogel, Wayne
;  APPLICANT:  Chen, Zhi
;  TITLE OF INVENTION:  Human Cytokine Receptor
;  FILE REFERENCE:  99-93
;  CURRENT APPLICATION NUMBER: US/09/728,911
;  CURRENT FILING DATE:  2000-12-01
;  PRIOR APPLICATION NUMBER:  US 60/169,049
;  PRIOR FILING DATE:  1999-12-03
;  PRIOR APPLICATION NUMBER:  US 60/232,219
;  PRIOR FILING DATE:  2000-09-13
;  PRIOR APPLICATION NUMBER:  US 60/244,610
;  PRIOR FILING DATE:  2000-10-31
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 231
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-728-911-2

  Query Match             5.8%;  Score 77.5;  DB 10;  Length 231;
  Best Local Similarity   25.8%;  Pred. No. 3.4;
  Matches   34;  Conservative   14;  Mismatches   41;  Indels   43;  Gaps    7;

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQ----AFHTYGLEWTPN----------YVRW 112
           |   |  :|| || | |   | :: |   |     ||   |:| |           :|::
Db       9 GFLISFFLTGVAGTQST---HESLKPQRVQFQSRNFHNI-LQWQPGRALTGNSSVYFVQY 64

Qy     113 TVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSES----------AAWVGQFDESKLPLFQ 161
           : ||   | :     :  ||| |  :| |  |           ||  | : |
Db      65 KIYGQRQWKNK----EDCWGTQELSCDLTSETSDIQEPYYGRVRAASAGSYSE------- 113

Qy     162 FINWVKVYKYTP 173
               |    ::||
```

```
Db    114 ---WSMTPRFTP 122


RESULT 15
US-09-949-192-6
; Sequence 6, Application US/09949192
; Patent No. US20020142292A1
; GENERAL INFORMATION:
;  APPLICANT: Parham, Christi L.
;  APPLICANT:  Gorman, Daniel L.
;  APPLICANT:  Kurata, Hirokazu
;  APPLICANT:  Arai, Naoko
;  APPLICANT:  Sana, Theodore R.
;  APPLICANT:  Mattson, Jeanine D.
;  APPLICANT:  Murphy, Erin E.
;  APPLICANT:  Savkoor, Chetan
;  APPLICANT:  Grein, Jeffery
;  APPLICANT:  Smith, Kathleen M.
;  APPLICANT:  McClanahan, Terrill K.
;  TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS AND METHODS
;  FILE REFERENCE: DX01169K
;  CURRENT APPLICATION NUMBER: US/09/949,192
;  CURRENT FILING DATE:  2001-09-07
;  PRIOR APPLICATION NUMBER: 60/231,267
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 231
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-949-192-6

  Query Match             5.8%;  Score 77.5;  DB 10;  Length 231;
  Best Local Similarity   25.8%;  Pred. No. 3.4;
  Matches   34;  Conservative   14;  Mismatches   41;  Indels   43;  Gaps    7;

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQ----AFHTYGLEWTPN----------YVRW 112
           |   |  :|| || | |   | :: |   |     ||   |:| |           :|::
Db       9 GFLISFFLTGVAGTQST---HESLKPQRVQFQSRNFHNI-LQWQPGRALTGNSSVYFVQY 64

Qy     113 TVDGQEVRKTEGGQVSNLTGTQGLRFNLWSSES----------AAWVGQFDESKLPLFQ 161
             : ||   | :     :  ||| |  :| |  |           ||  | : |
Db      65 KIYGQRQWKNK----EDCWGTQELSCDLTSETSDIQEPYYGRVRAASAGSYSE------- 113

Qy     162 FINWVKVYKYTP 173
             |    ::||
Db     114 ---WSMTPRFTP 122


Search completed: January  9, 2003, 12:14:39
Job time : 7.31481 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:57 ; Search time 25.8333 Seconds
(without alignments)
1279.206 Million cell updates/sec

Title: US-09-654-652A-1
Perfect score: 1333
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....TRKGQESFNGQVPRDDEPAP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	257.5	19.3	308	20	AAW93001	B. alkalophilus be
2	246.5	18.5	242	10	AAP95000	Bacillus subtilis
3	244	18.3	214	22	AAE07317	Barley recombinant
4	244	18.3	239	11	AAR06621	Hybrid (1,3-1,4)-p
5	226	17.0	234	11	AAR03775	Thermostable beta-
6	226	17.0	237	11	AAR05803	Heat-stable endo-b
7	225	16.9	208	23	ABB76858	Bacterial lichenas
8	225	16.9	214	23	ABB76859	Bacterial lichenas
9	220.5	16.5	237	11	AAR06622	Hybrid (1,3-1,4)-p
10	189	14.2	245	19	AAW37884	Lichenase protein.
11	184	13.8	282	21	AAG48467	Arabidopsis thalia
12	184	13.8	282	23	ABB93737	Herbicidally activ
13	172	12.9	282	21	AAG32464	Arabidopsis thalia
14	172	12.9	282	23	ABB91456	Herbicidally activ
15	171.5	12.9	277	21	AAG52114	Arabidopsis thalia
16	171.5	12.9	277	23	ABB93169	Herbicidally activ
17	169.5	12.7	269	21	AAG18644	Arabidopsis thalia
18	169.5	12.7	282	21	AAG18643	Arabidopsis thalia
19	169.5	12.7	282	23	ABB93168	Herbicidally activ
20	169.5	12.7	298	21	AAG18642	Arabidopsis thalia
21	163	12.2	286	23	ABB93103	Herbicidally activ
22	161	12.1	284	23	ABB93853	Herbicidally activ
23	159	11.9	287	23	ABB93104	Herbicidally activ
24	154.5	11.6	284	23	ABB93854	Herbicidally activ
25	154	11.6	845	22	AAB98062	Clostridium perfri
26	150.5	11.3	269	21	AAG20347	Arabidopsis thalia
27	150.5	11.3	269	21	AAG20350	Arabidopsis thalia
28	150.5	11.3	269	21	AAG52653	Arabidopsis thalia
29	150.5	11.3	269	23	ABB93167	Herbicidally activ
30	150.5	11.3	290	21	AAG20346	Arabidopsis thalia
31	150.5	11.3	297	21	AAG52652	Arabidopsis thalia
32	150.5	11.3	300	21	AAG20349	Arabidopsis thalia
33	149.5	11.2	269	21	AAG20329	Arabidopsis thalia
34	149.5	11.2	299	21	AAG20328	Arabidopsis thalia
35	149.5	11.2	306	17	AAR97362	Oerskovia beta-1,3
36	147	11.0	247	18	AAW11593	Protein encoded by
37	147	11.0	247	20	AAY08308	A. thaliana merist
38	145.5	10.9	285	23	ABB93852	Herbicidally activ
39	145	10.9	263	18	AAW29455	Oerskovia xanthine
40	143	10.7	204	21	AAG48468	Arabidopsis thalia
41	143	10.7	435	18	AAW29456	Oerskovia xanthine
42	141	10.6	261	12	AAR11599	Beta-1,3-glucanase
43	140.5	10.5	303	18	AAW29457	Oerskovia xanthine
44	140.5	10.5	307	21	AAG40361	Arabidopsis thalia
45	140.5	10.5	310	21	AAG40360	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW93001
ID AAW93001 standard; Protein; 308 AA.
XX
AC AAW93001;
XX
DT 19-MAY-1999 (first entry)
XX
DE B. alkalophilus beta-glucanase protein.
XX
KW Beta-glucanase; enzyme; glucan; removal; lichenan; membrane;
KW equipment; food industry; brewing.
XX
OS Bacillus alkalophilus.
XX
PN DE19732751-A1.
XX
PD 04-FEB-1999.
XX
PF 30-JUL-1997; 97DE-1032751.
XX
PR 30-JUL-1997; 97DE-1032751.
XX
PA (HENK) HENKEL KGAA.
XX
PI Hillen W, Maurer K;
XX
DR WPI; 1999-122161/11.
DR N-PSDB; AAX02912.
XX
PT New Bacillus alkalophilus beta-glucanase protein and gene - useful
PT for removing glucan and/or lichenan from membranes in the brewing
PT industry

```
XX
PS   Claim 6; Page 5; 8pp; German.
XX
CC   This sequence represents a novel beta-glucanase isolated from Bacillus
CC   alkalophilus DSM 9956. This enzyme is useful for removing glucan and/or
CC   lichenan from membranes and equipment in the food industry, especially
CC   the brewing industry.
XX
SQ   Sequence   308 AA;

  Query Match             19.3%;  Score 257.5;  DB 20;  Length 308;
  Best Local Similarity   36.5%;  Pred. No. 4.5e-17;
  Matches   62;  Conservative   22;  Mismatches   73;  Indels   13;  Gaps    6;

Qy       6 DFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN 65
           ::   || | :  ||| ||  || | ::||||||:| |  |    |  || |:||| |||:
Db      96 EYKAGELRTNQFYQYGLFEVNMKPAKSTGTVSSLFTY-TGPWDWDNDPWDEIDIEFLGKD 154

Qy      66 PGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGG 125
             | |  |   |    :| :| :   | ::|:||  || |  : | |:|: |
Db     155 TTRVQFNYFTNGVG---NNEHYHELGFDASESFNTYAFEWRPESISWYVNGELVYTA--- 208

Qy     126 QVSNLTGT-QGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYTP 173
             |:  | | :  |||       | | ||    |:   :||   :|||
Db     209 -TENIPQTPQKIMMNLWPGIGVDGWTGVFDGEDTPVVTEYDWV---RYTP 254


RESULT 2
AAP95000
ID   AAP95000 standard; Protein; 242 AA.
XX
AC   AAP95000;
XX
DT   12-FEB-1997  (first entry)
XX
DE   Bacillus subtilis lichenase.
XX
KW   Beer production; fermentation; barley; beta-glucan; hydrolysis;
KW   lichenase.
XX
OS   Bacillus subtilis Y-25.
XX
PN   JP01067181-A.
XX
PD   13-MAR-1989.
XX
PF   08-SEP-1987;   87JP-0224615.
XX
PR   08-SEP-1987;   87JP-0224615.
XX
PA   (ASAK ) ASAHI BREWERIES KK.
XX
DR   WPI; 1989-119863/16.
DR   N-PSDB; AAN95000.
XX
PT   Recombinant plasmid used in beer prodn. - obtd. by integrating
PT   lichenase gene derived from Bacillus subtilis, into vector
XX
PS   Disclosure; Fig 4; 7pp; Japanese.
XX
CC   The lichenase gene from Bacillus subtilis Y-25 is used for
CC   transforming Bacillus hosts so that they show increased lichenase
CC   expression. The recombinant lichenase enzyme produced by the
CC   transformants is useful in beer production for decomposing beta-
CC   glucan from barley.
XX
SQ   Sequence   242 AA;

  Query Match             18.5%;  Score 246.5;  DB 10;  Length 242;
  Best Local Similarity   35.5%;  Pred. No. 4e-16;
  Matches   60;  Conservative   24;  Mismatches   74;  Indels   11;  Gaps    6;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :   || || |:||| |
Db      83 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPT---DGTPWDEIDIEFL 139

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |
Db     140 GKDTTKVQFNYYTNGAG---NHEKIVDLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 195

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
             |:    |   :  |||:      |:| ::    ||:   :||:  |
Db     196 ATNQIPTTPGK--IMMNLWNGTGVDEWLGSYNGVN-PLYAHYDWVRYTK 241


RESULT 3
AAE07317
ID   AAE07317 standard; Protein; 214 AA.
XX
AC   AAE07317;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Barley recombinant, thermostable (1,3-1,4)-beta-glucanase enzyme.
XX
KW   Barley; transgenic barley malt; carbohydrate degrading enzyme; chicken;
KW   (1,3-1,4)-beta-glucanase; nutritional value; foodstuff; thermostable.
XX
OS   Hordeum vulgare.
XX
PN   WO200159141-A2.
XX
PD   16-AUG-2001.
XX
PF   09-FEB-2001; 2001WO-US04222.
XX
PR   10-FEB-2000; 2000US-0181473.
PR   09-NOV-2000; 2000US-0247126.
XX
PA   (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI   Von Wettstein D,  Huang J,  Horvath H;
XX
DR   WPI; 2001-497082/54.
XX
PT   New barley based foodstuff for animals, i.e. chicken comprising
PT   recombinant carbohydrate degrading enzyme (1,3-1,4)-beta-glucanase -
XX
PS   Claim 8; Page 37-38; 43pp; English.
XX
CC   The present invention relates to a food stuff comprising barley feed and
CC   transgenic barley malt where the transgenic barley malt comprises a
CC   recombinant carbohydrate degrading enzyme comprising a (1,3-1,4)-beta-
CC   glucanase. The food stuff is useful to feed animal, preferably chickens
CC   Barley is cheaper than corn, the principal foodstuff of chickens.
CC   Chickens cannot efficiently utilise barley as an energy source as they
CC   do not possess a gut enzyme that depolymerises beta-D-glucan, major
CC   carbohydrate present in the barley endosperm. The invention provides a
CC   barley based foodstuff which comprises a recombinant carbohydrate
CC   degrading enzyme improving the nutritional value of the foodstuff.
CC   The present sequence is barley recombinant, thermostable
CC   (1,3-1,4)-beta-glucanase enzyme.
XX
SQ   Sequence   214 AA;

  Query Match             18.3%;  Score 244;  DB 22;  Length 214;
  Best Local Similarity   35.5%;  Pred. No. 5.9e-16;
  Matches   61;  Conservative   26;  Mismatches   69;  Indels   16;  Gaps    8;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      55 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 111

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:    | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db     112 GKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT 167
```

```
Qy    123 EGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
              :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db    168 ---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 212
```

RESULT 4
AAR06621
ID AAR06621 standard; protein; 239 AA.
XX
AC AAR06621;
XX
DT 09-JAN-1991 (first entry)
XX
DE Hybrid (1,3-1,4)-pre-beta-glucanase.
XX
KW Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.
XX
OS Bacillus amyloliquefaciens, Bacillus macerans.
XX
FH Key Location/Qualifiers
FT Domain 1..129
FT /label=amino terminal of beta-amyloliquefaciens
FT Domain 133..236
FT /label=carboxyl-terminal of B.macerans
XX
PN WO9009436-A.
XX
PD 23-AUG-1990.
XX
PF 16-FEB-1990; 90WO-DK00044.
XX
PR 04-AUG-1989; 89DK-0003848.
PR 16-FEB-1989; 89DD-0325800.
XX
PA (CARL-) CARLSBERG A/S.
PA (DEAK) AKAD WISSENSCHAFT DDR.
XX
PI Borriss R, Hofemeister J, Thomsen KK, Olsen O, Vonwettstein D;
XX
DR WPI; 1990-275129/36.
DR N-PSDB; AAQ05832.
XX
PT New thermostable (1,3-1,4)-beta-glucanase - prepd. using hybrid
PT gene obtd. using Bacillus amyloliquefaciens and B.macerans genes
XX
PS Disclosure; page 26; 84pp; English.
XX
CC This hybrid protein is encoded by the beta-glucanase-H1 gene.
CC Following processing of the signal peptide the mature protein
CC is produced, comprising the amino terminus of the amyloliquefaciens beta-glucanase and the carboxyl-terminal half of the
CC B.macerans beta-glucanase. This hybrid protein is thermostable
CC and hydrolyses beta-glycosidic linkages in (1,3-1,4)-beta-glucans.
CC Reducing sugars are obtd. at high temps. and thus this enzyme can
CC be used in the mfr. of food prods., esp. beer and animal feed (eg
CC for feeding poultry). See also AAQ05833.
XX
SQ Sequence 239 AA;

Query Match 18.3%; Score 244; DB 11; Length 239;
Best Local Similarity 35.5%; Pred. No. 6.9e-16;
Matches 61; Conservative 26; Mismatches 69; Indels 16; Gaps 8;

```
Qy      3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
          |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db     80 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 136

Qy     63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
          ||:   | |  |   |    ||  ::   | : |||| :| | |::| |||  :: |
Db    137 GKDTTKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT 192

Qy    123 EGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
```

```
              :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db    193 ---ATANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 237
```

RESULT 5
AAR03775
ID AAR03775 standard; protein; 234 AA.
XX
AC AAR03775;
XX
DT 31-JUL-1990 (first entry)
XX
DE Thermostable beta-glucanase.
XX
KW Thermostable beta-glucanase; ss; Bacillus subtilis.
XX
OS Bacillus macerans.
XX
PN DD272102-A.
XX
PD 27-SEP-1989.
XX
PF 12-MAY-1988; 88DD-0315706.
XX
PR 12-MAY-1988; 88DD-0315706.
XX
PA (DEAK) AKAD WISSENSCHAFT DDR.
XX
PI Borriss R;
XX
DR WPI; 1990-067913/10.
DR N-PSDB; AAQ03519.
XX
PT Thermostable beta-glucanase production -
PT using Bacillus subtilis transformed with gene from Bacillus
PT macerans.
XX
PS Fig 1; ; 9pp; German.
XX
CC The gene encoding Bacillus macerans thermostable
CC beta-glucanase is expressed in Bacillus subtilis. The
CC enzyme is useful for lowering the viscosity of brewing
CC mashes and in the production of feedstuff.
XX
SQ Sequence 234 AA;

Query Match 17.0%; Score 226; DB 11; Length 234;
Best Local Similarity 35.1%; Pred. No. 4.1e-14;
Matches 59; Conservative 23; Mismatches 70; Indels 16; Gaps 8;

```
Qy      7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
          |  ||  :    || :|  || |  :| ||| | |     | |  | |:||| |||:
Db     79 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 135

Qy     67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
            | |  |   |     ||  ::   | : ||||  :| | |::| |||  :: |
Db    136 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 188

Qy    127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
           :|:  | | :  |||:      |:| :: :  ||:   :||   |||
Db    189 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 232
```

RESULT 6
AAR05803
ID AAR05803 standard; protein; 237 AA.
XX
AC AAR05803;
XX
DT 08-NOV-1990 (first entry)
XX
DE Heat-stable endo-beta-1,3-1,4-glucanase.
XX

KW Endo-beta-1,3-1,4-glucanase; barley; brewing.
XX
OS Bacillus macerans.
XX
PN DD275704-A.
XX
PD 31-JAN-1990.
XX
PF 23-SEP-1988; 88DD-0320082.
XX
PR 23-SEP-1988; 88DD-0320082.
XX
PA (DEAK) AKAD WISSENSCHAFT DDR.
XX
PI Borriss R, Wobus U, Mendel R-R, Baumlein H;
XX
DR WPI; 1990-210631/28.
DR N-PSDB; AAQ05167.
XX
PT Prepn. of barley plants expressing heat stable beta-glucanase -
PT by transforming cells with appropriate vector then regeneration
PT giving seeds useful in brewing without conversion to malt.
XX
PS Disclosure; ; p; German.
XX
CC The corresp. gene and the plant promoter region are inserted into an
CC expression vector and used to transform barley cells. The transformants
CC are used to regenerate barley plants which are useful in brewing. This
CC protein accumulates in the ripe (but ungerminated) seeds. The
CC proportion of these to malted seeds is increased therefore, without
CC an unacceptable increase in viscosity.
XX
SQ Sequence 237 AA;

Query Match 17.0%; Score 226; DB 11; Length 237;
Best Local Similarity 35.1%; Pred. No. 4.2e-14;
Matches 59; Conservative 23; Mismatches 70; Indels 16; Gaps 8;

```
Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db      82 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
Db     139 TKVQFNYYTNGVGGH---EKVISLGFDASKGFHTYAFDWQPGYIKWYVDG-VLKHT---A 191

Qy     127 VSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172
Db     192 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 235
```

RESULT 7
ABB76858
ID ABB76858 standard; Protein; 208 AA.
XX
AC ABB76858;
XX
DT 27-JUN-2002 (first entry)
XX
DE Bacterial lichenase #1.
XX
KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.
XX
OS Unidentified.
XX
PN WO200212511-A1.
XX
PD 14-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-ES00303.
XX
PR 28-JUL-2000; 2000ES-0001922.
XX
PA (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX
PI Rodriguez Munoz V, Perez Mellado R;
XX
DR WPI; 2002-217195/27.
DR N-PSDB; ABL53374.
XX
PT Producing nucleic acid encoding lichenase, useful for producing enzymes
PT for improving filtration in brewing, comprises isolating the nucleic
PT acid from soil bacteria -
XX
PS Claim 18; Page 18-19; 27pp; Spanish.
XX
CC The present sequence is a protein sequence from a soil bacterium for an
CC enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)
CC activity. The enzyme is useful in brewing and for degrading
CC beta-glucanases that cause problems during filtration.
XX
SQ Sequence 208 AA;

Query Match 16.9%; Score 225; DB 23; Length 208;
Best Local Similarity 36.2%; Pred. No. 4.3e-14;
Matches 55; Conservative 21; Mismatches 66; Indels 10; Gaps 5;

```
Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db      58 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 114

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
Db     115 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 170

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 153
Db     171 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYN 200
```

RESULT 8
ABB76859
ID ABB76859 standard; Protein; 214 AA.
XX
AC ABB76859;
XX
DT 27-JUN-2002 (first entry)
XX
DE Bacterial lichenase #2.
XX
KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.
XX
OS Unidentified.
XX
PN WO200212511-A1.
XX
PD 14-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-ES00303.
XX
PR 28-JUL-2000; 2000ES-0001922.
XX
PA (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX
PI Rodriguez Munoz V, Perez Mellado R;
XX
DR WPI; 2002-217195/27.
DR N-PSDB; ABL53375.
XX
PT Producing nucleic acid encoding lichenase, useful for producing enzymes
PT for improving filtration in brewing, comprises isolating the nucleic
PT acid from soil bacteria -
XX
PS Claim 18; Page 21-22; 27pp; Spanish.
XX
CC The present sequence is a protein sequence from a soil bacterium for an

```
CC   enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)
CC   activity. The enzyme is useful in brewing and for degrading
CC   beta-glucanases that cause problems during filtration.
XX
SQ   Sequence   214 AA;

  Query Match             16.9%;  Score 225;  DB 23;  Length 214;
  Best Local Similarity   36.2%;  Pred. No. 4.5e-14;
  Matches   55;  Conservative   21;  Mismatches   66;  Indels   10;  Gaps    5;

Qy       3 SAKDFSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
           |   |   |  :::   || :| ||| |  :| ||| | |   :|   | || |:||| |
Db      63 SYNKFDCGENRSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 119

Qy      63 GKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
           ||:   | |  |  ||    ||   :   |  |:||| :| || ::| |||| :: |
Db     120 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHT 175

Qy     123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 153
             |:   |  :  |||:     |:| ::
Db     176 ATTQIPAAPGK--IMMNLWNGTGVDDWLGSYN 205


RESULT 9
AAR06622
ID   AAR06622 standard; protein; 237 AA.
XX
AC   AAR06622;
XX
DT   09-JAN-1991  (first entry)
XX
DE   Hybrid (1,3-1,4)-pre-beta-glucanase.
XX
KW   Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.
XX
OS   Bacillus amyloliquefaciens, Bacillus macerans.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..127
FT                   /label=amino-terminal half of B.macerans beta-glucanase
FT   Domain          131..234
FT                   /label=carboxyl-terminal half of B.amyloliquefaciens
FT                         beta-glucanase
XX
PN   WO9009436-A.
XX
PD   23-AUG-1990.
XX
PF   16-FEB-1990;   90WO-DK00044.
XX
PR   04-AUG-1989;   89DK-0003848.
PR   16-FEB-1989;   89DD-0325800.
XX
PA   (CARL-) CARLSBERG A/S.
PA   (DEAK ) AKAD WISSENSCHAFT DDR.
XX
PI   Borriss R,  Hofemeister J,  Thomsen KK,  Olsen O,  Vonwettstein D;
XX
DR   WPI; 1990-275129/36.
DR   N-PSDB; AAQ05833.
XX
PT   New thermostable (1,3-1,4)-beta-glucanase - prepd. using hybrid
PT   gene obtd. using Bacillus amyloliquefaciens and B.macerans genes
XX
PS   Disclosure; page 28; 84pp; English.
XX
CC   This hybrid protein is encoded by the beta-glucanase-H2 gene.
CC   Following processing of the signal peptide the mature protein
CC   is produced, comprising the amino terminus of the B.macerans
CC   beta-glucanase and the carboxyl-terminal half of the B.amylo-
CC   liquefaciens beta-glucanase.  This hybrid protein is thermo-
CC   stable and hydrolyses beta-glycosidic linkages in (1,3-1,4)-
CC   beta-glucans.  Reducing sugars are obtd. at high temps. and
CC   thus this enzyme can be used in the mfr. of food prods., esp.
CC   beer and animal feed (e.g.for feeding poultry).  See also
CC   AAQ05833.
XX
SQ   Sequence   237 AA;

  Query Match             16.5%;  Score 220.5;  DB 11;  Length 237;
  Best Local Similarity   33.9%;  Pred. No. 1.5e-13;
  Matches   56;  Conservative   23;  Mismatches   75;  Indels   11;  Gaps    6;

Qy       7 FSGAELYTLEEVQYGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
           |  ||  :     || :|  || |  :| ||| | |       | |  | |:||| |||:
Db      82 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138

Qy      67 GSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQ 126
             | |  |  ||     ||   :   |  |:|||  :| || ::| |||| :: |   |
Db     139 TKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDWQPNSIKWYVDGQ-LKHTATTQ 194

Qy     127 VSNLTGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYK 170
           :    |   :  |||:      |:| ::    |::   :|::  |
Db     195 IPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYHWMRYRK 236


RESULT 10
AAW37884
ID   AAW37884 standard; Protein; 245 AA.
XX
AC   AAW37884;
XX
DT   20-AUG-1998  (first entry)
XX
DE   Lichenase protein.
XX
KW   Lichenase; licA; fungus; enzyme; beta-1,4-glucan bond hydrolysation;
KW   beta-1,3-linked glucan bond; grain-containing feed; grain treatment;
KW   nutrient availability; brewing; fermentation.
XX
OS   Orpinomyces sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..29
FT                   /note= "signal peptide"
FT   Protein         30..246
FT                   /note= "mature lichenase"
XX
PN   WO9814595-A1.
XX
PD   09-APR-1998.
XX
PF   03-OCT-1997;   97WO-US17811.
XX
PR   04-OCT-1996;   96US-0027882.
XX
PA   (UYGE-) UNIV GEORGIA RES FOUND INC.
XX
PI   Chen H,  Li X,  Ljungdahl LG;
XX
DR   WPI; 1998-240094/21.
DR   N-PSDB; AAV29067.
XX
PT   New isolated lichenase protein - is obtained from Orpinomyces PC-2,
PT   used for treatment of grain to improve feeds or to improve brewing
PT   and fermentation processes
XX
PS   Claim 1; Page 24-25; 41pp; English.
XX
CC   This sequence is the lichenase (licA) of Orpinomyces sp. strain PC-2
CC   of the invention. The protein was purified from a fungus or a fungal
CC   culture or from a recombinant DNA molecule having a fungal lichenase
CC   coding sequence. The lichenase enzyme hydrolyses the beta-1,4-glucan
CC   bonds adjacent to beta-1,3-linked glucan bonds, but does not cleave
CC   beta-1,4-linked glucans. They can be used for the treatment of animal
CC   grain-containing feeds to improve nutrient availability and for treatment
```

```
CC   of grain (e.g. barley or wheat) in the brewing and fermentation
CC   industries to increase carbon substrate availability and to maximise
CC   production of desired products.
XX
SQ   Sequence   245 AA;

  Query Match            14.2%;  Score 189;  DB 19;  Length 245;
  Best Local Similarity  32.9%;  Pred. No. 2e-10;
  Matches   56;  Conservative   16;  Mismatches   76;  Indels   22;  Gaps    8;

Qy       5 KDFSGAELYTLEEVQ------YGKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVD 58
           :| ||   ||  | :      || |:  ||     | ||| | |   |   ||  | |:|
Db      83 RDGSG---YTCGEYRTKNYYGYGMFQVNMKPIKNPGVVSSFFTYTGPS---DGTKWDEID 136

Qy      59 IEVLGKNPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQE 118
           || || :    | |  |   |     |  | :   | | |||||  |  | : | |||
Db     137 IEFLGYDTTKVQFNYYTNGQGHH---EHIHYLGFDASQGFHTYGFFWARNSITWYVDGTA 193

Qy     119 VRKTEGGQVSNLTGTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWV 166
           |         |:  | | :  | |:      |:  |:  :  :  : :||
Db     194 VYTA----YDNIPDTPGKIMMNAWNGIGVDDWLRPFN-GRTNISAYYDWV 238


RESULT 11
AAG48467
ID   AAG48467 standard; Protein; 282 AA.
XX
AC   AAG48467;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 61207.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.8%; Score 184; DB 21; Length 282;
Best Local Similarity 27.6%; Pred. No. 7.8e-10;
Matches 59; Conservative 33; Mismatches 98; Indels 24; Gaps 9;

```
Qy       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDI 59
           :|   |||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:|
Db      54 LSLDKFSGSGFQSHQEFLYGKVEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDF 107

Qy      60 EVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 116
           | ||    :| :  :|: | |    |  : |    |  |  |||| : | |  | :|:||
Db     108 EFLGNISGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVD--FHTYCIIWNPQRVIFTIDG 164

Qy     117 QEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKVY 169
             :|: :  :   :     | :|   :|| :|  |  |   : | || |   |     |
Db     165 IPIREFKNSEALGVPFPKHQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAFYRNYNVD 224

Qy     170 KYTPGQGEGGSDFTLDWTDNFDTFDGS---RWGK 200
             |:      |      | |    :| :
Db     225 ACVWSNGKSSCSANSSWFTQVLDFKGKNRVKWAQ 258
```

RESULT 12
ABB93737
ID ABB93737 standard; Protein; 282 AA.
XX
AC ABB93737;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2948.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 2948; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX

```
SQ   Sequence   282 AA;

  Query Match             13.8%;  Score 184;  DB 23;  Length 282;
  Best Local Similarity   27.6%;  Pred. No. 7.8e-10;
  Matches   59;  Conservative   33;  Mismatches   98;  Indels   24;  Gaps    9;

Qy       2 VSAKDFSGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDI 59
           :|   |||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:|
Db      54 LSLDKFSGSGFQSHQEFLYGKVEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDF 107

Qy      60 EVLGK---NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDG 116
           | ||    :| :  :|: | |    |  : |    |  | ||| : | |  | :|:||
Db     108 EFLGNISGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVD--FHTYCIIWNPQRVIFTIDG 164

Qy     117 QEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKVY 169
             :|: :  :   :     | :|   :|| :|  |  |   : | || |   |     |
Db     165 IPIREFKNSEALGVPFPKHQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAFYRNYNVD 224

Qy     170 KYTPGQGEGGSDFTLDWTDNFDTFDGS---RWGK 200
                |:        |      | |    :| :
Db     225 ACVWSNGKSSCSANSSWFTQVLDFKGKNRVKWAQ 258


RESULT 13
AAG32464
ID   AAG32464 standard; Protein; 282 AA.
XX
AC   AAG32464;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 39167.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 12.9%; Score 172; DB 21; Length 282;
Best Local Similarity 25.9%; Pred. No. 1.2e-08;
Matches 63; Conservative 34; Mismatches 88; Indels 58; Gaps 12;

```
Qy       8 SGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK- 64
Db      60 SGSGFQSNQEFLYGKAEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDFEFLGNI 113

Qy      65 --NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
Db     114 SGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVN--FHTYCITWNPQRIIFTVDGIPIREF 170

Qy     123 EGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKVYKYTPGQ 175
Db     171 KNPEAIGVPFPTRQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAF------YR----- 219

Qy     176 GEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQES 235
Db     220 ------------NYNV-DGCVWANGKSSCSANSPWFTQK-----------LDSNGQTR 253

Qy     236 FNG 238
Db     254 MKG 256
```

RESULT 14
ABB91456
ID ABB91456 standard; Protein; 282 AA.
XX
AC ABB91456;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 667.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 667; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant

```
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   282 AA;

  Query Match             12.9%;  Score 172;  DB 23;  Length 282;
  Best Local Similarity   25.9%;  Pred. No. 1.2e-08;
  Matches   63;  Conservative   34;  Mismatches   88;  Indels   58;  Gaps   12;

Qy       8 SGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK- 64
           ||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:| | ||
Db      60 SGSGFQSNQEFLYGKAEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDFEFLGNI 113

Qy      65 --NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
             :| :  :|: | |    |  : |    |  :  |||| : | |  : :|||| :|:
Db     114 SGHPYTLHTNVYT-KGTGDKEQQFHLWFDPTVN--FHTYCITWNPQRIIFTVDGIPIREF 170

Qy     123 EGGQVSNL--TGTQGLRF--NLWSSESAAWVG---QFDESKLPLFQFINWVKVYKYTPGQ 175
           :  :   :     | :|   :|| :|  |  |   : | || |   |      |:
Db     171 KNPEAIGVPFPTRQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAF------YR----- 219

Qy     176 GEGGSDFTLDWTDNFDTFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQES 235
                      |::  ||  |  |  :   |    | |           |   ||
Db     220 ------------NYNV-DGCVWANGKSSCSANSPWFTQK-----------LDSNGQTR 253

Qy     236 FNG 238
             |
Db     254 MKG 256


RESULT 15
AAG52114
ID   AAG52114 standard; Protein; 277 AA.
XX
AC   AAG52114;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 66212.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
```

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 12.9%; Score 171.5; DB 21; Length 277;
Best Local Similarity 27.0%; Pred. No. 1.3e-08;
Matches 53; Conservative 32; Mismatches 70; Indels 41; Gaps 9;

```
Qy       8 SGAELYTLEEVQYGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK- 64
           ||:   : :|  ||| | :||:    ::|||:: :|       : |  | |:| | ||
Db      55 SGSGFQSNQEFLYGKAEVQMKLVPGNSAGTVTTFYLK------SPGTTWDEIDFEFLGNI 108

Qy      65 --NPGSFQSNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKT 122
             :| :  :|: | |    |  : |    | |:  |||| : | |  : :||||  :|:
Db     109 SGHPYTLHTNVYT-KGSGDKEQQFHLWFDPTAN--FHTYCITWNPQRIIFTVDGIPIREF 165

Qy     123 EGGQVSNLTGTQGLRF----------NLWSSESAAWVG---QFDESKLPLFQFIN----- 164
             :      ::|: |          :|| :|  |  |   : | || |   :
Db     166 MNAE------SRGVPFPTKQPMRLYASLWEAEHWATRGGLEKTDWSKAPFTAYYRNYNVE 219

Qy     165 ---WVKVYKYTPGQGE 177
               ||      |   :
Db     220 GCVWVNGKSVCPANSQ 235
```

Search completed: January 9, 2003, 12:16:17
Job time : 27.8333 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:04:12 ; Search time 11.7431 Seconds
(without alignments)
1613.076 Million cell updates/sec

Title: US-09-654-652A-2
Perfect score: 1439
Sequence: 1 MVSAKDFSGAELYTLEEVQY.....PNSSSVDKLAAALEHHHHHH 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 263044 seqs, 70945482 residues

Total number of hits satisfying chosen parameters: 263044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	116.5	8.1	673	6	US-10-294-561-3	Sequence 3, Appli
2	113	7.9	21	6	US-10-010-160-68	Sequence 68, Appl
3	104.5	7.3	1510	5	US-09-724-676-83340	Sequence 83340, A
4	104.5	7.3	1510	5	US-09-724-676A-83340	Sequence 83340, A
5	103.5	7.2	1554	5	US-09-724-676-83341	Sequence 83341, A
6	103.5	7.2	1554	5	US-09-724-676A-83341	Sequence 83341, A
7	100.5	7.0	526	5	US-09-455-294A-10	Sequence 10, Appl
8	99.5	6.9	1558	5	US-09-724-676-83343	Sequence 83343, A
9	99.5	6.9	1558	5	US-09-724-676A-83343	Sequence 83343, A
10	98.5	6.8	1602	5	US-09-724-676-83344	Sequence 83344, A
11	98.5	6.8	1602	5	US-09-724-676A-83344	Sequence 83344, A
12	97.5	6.8	1629	5	US-09-724-676-83342	Sequence 83342, A
13	97.5	6.8	1629	5	US-09-724-676A-83342	Sequence 83342, A
14	97	6.7	481	6	US-10-141-531-59	Sequence 59, Appl
15	94	6.5	290	5	US-09-424-705B-4	Sequence 4, Appli
16	92	6.4	381	6	US-10-141-531-47	Sequence 47, Appl
17	92	6.4	381	6	US-10-141-531-48	Sequence 48, Appl
18	92	6.4	381	6	US-10-141-531-49	Sequence 49, Appl
19	92	6.4	381	6	US-10-141-531-50	Sequence 50, Appl
20	92	6.4	381	6	US-10-141-531-51	Sequence 51, Appl
21	92	6.4	381	6	US-10-141-531-52	Sequence 52, Appl
22	92	6.4	381	6	US-10-141-531-53	Sequence 53, Appl
23	92	6.4	381	6	US-10-141-531-54	Sequence 54, Appl
24	92	6.4	381	6	US-10-141-531-56	Sequence 56, Appl
25	92	6.4	381	6	US-10-141-531-58	Sequence 58, Appl
26	90	6.3	378	6	US-10-114-104A-122	Sequence 122, App
27	90	6.3	398	6	US-10-114-104A-121	Sequence 121, App
28	86.5	6.0	438	6	US-10-276-774-2151	Sequence 2151, Ap
29	86.5	6.0	444	6	US-10-092-411A-4346	Sequence 4346, Ap
30	83.5	5.8	486	5	US-09-724-676-60984	Sequence 60984, A
31	83.5	5.8	486	5	US-09-724-676A-60984	Sequence 60984, A
32	83.5	5.8	638	5	US-09-917-384A-9	Sequence 9, Appli
33	83.5	5.8	640	5	US-09-917-384A-5	Sequence 5, Appli
34	83.5	5.8	1121	5	US-09-917-384A-1	Sequence 1, Appli
35	83.5	5.8	1125	6	US-10-276-781-1219	Sequence 1219, Ap
36	82.5	5.7	261	1	PCT-US02-33165-3	Sequence 3, Appli
37	81	5.6	246	1	PCT-US02-36496-1920	Sequence 1920, Ap
38	81	5.6	246	6	US-10-293-418-1920	Sequence 1920, Ap
39	80.5	5.6	246	1	PCT-US02-36496-1264	Sequence 1264, Ap
40	80.5	5.6	246	6	US-10-293-418-1264	Sequence 1264, Ap
41	80.5	5.6	250	1	PCT-US02-36496-1280	Sequence 1280, Ap
42	80.5	5.6	250	6	US-10-293-418-1280	Sequence 1280, Ap
43	80	5.6	634	5	US-09-455-294A-8	Sequence 8, Appli
44	78.5	5.5	209	6	US-10-203-138A-14093	Sequence 14093, A
45	78.5	5.5	248	1	PCT-US02-36496-861	Sequence 861, App

ALIGNMENTS

RESULT 1
US-10-294-561-3
; Sequence 3, Application US/10294561
; GENERAL INFORMATION:
; APPLICANT: Seikagaku Corporation
; TITLE OF INVENTION: (1-3)-Beta-D-Glucan Binding Domain Protein, Measuring Method U
; TITLE OF INVENTION: Substance And Assay Kit
; FILE REFERENCE: Q72918
; CURRENT APPLICATION NUMBER: US/10/294,561
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: JP 2001-351943
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Tachypleus tridentatus
US-10-294-561-3

Query Match 8.1%; Score 116.5; DB 6; Length 673;
Best Local Similarity 22.8%; Pred. No. 0.018;
Matches 61; Conservative 43; Mismatches 91; Indels 73; Gaps 16;

```
Qy       1 MVSAK--DFSGAELYTLEEV--------QYGKFEARMKMAAASGTVSSMFLYQNGSEIAD 50
           :::||  |: | : ||   :         :||| ||:| : :  | |  ||
Db      73 VITAKREDYDGFK-YTSARLKTQFDKSWKYGKIEAKMAIPSFRG-VWVMF---------- 120

Qy      51 GRPWVEVDIEVLGKNPGSFQSNIITGK-AGAQKTSEKHHAVSPAADQA------------ 97
             |:  |     : | | : : |  :    :|      |  :|    |
Db     121 ---WMSGDNTNYVRWPSSGEIDFIEHRNTNNEKVRGTIHWSTPDGAHAHHNRESNTNGID 177

Qy      98 FHTYGLEWTPNYVRWTVDGQEV--RKTEGGQVSNLTGTQGLRFNLWSSESAA----WVGQ 151
           :| | :||  : |:| |:| :    | :||    : |    |  ::   : |    | |
Db     178 YHIYSVEWNSSIVKWFVNGNQYFEVKIQGG----VNGKSAFRNKVFVILNMAIGGNWPG- 232

Qy     152 FD--ESKLPLFQFINWVKVYKYTPGQGEGG-------------SDFTLDWTDNFDTFDG 195
           ||  :   |   :|::|:||:        |              |:  || ||  :  ||
Db     233 FDVADEAFPAKMYIDYVRVYQDASTSSPVGDTSLDGYYFVQNRHSELYLDVTDASNE-DG 291

Qy     196 SRWGKGDWTFDGNR-----VDLTDKNIY 218
           :   :  |:: ||       :  : |:|
Db     292 AFLQQ--WSYSGNENQQFDFEHLENNVY 317
```

RESULT 2
US-10-010-160-68
; Sequence 68, Application US/10010160
; GENERAL INFORMATION:

```
;  APPLICANT: Rosey, Everett L.
;  APPLICANT:  Strugnell, Richard A.
;  APPLICANT:  Good, Robert T,
;  APPLICANT:  King, Kendall W.
;  TITLE OF INVENTION: NOVEL THERAPEUTIC COMPOSITIONS FOR
;  TITLE OF INVENTION:  TREATING INFECTION BY LAWSONIA SPP.
;  FILE REFERENCE: DAVI110.001AUS
;  CURRENT APPLICATION NUMBER: US/10/010,160
;  CURRENT FILING DATE:  2001-11-09
;  PRIOR APPLICATION NUMBER: AU PR1381
;  PRIOR FILING DATE: 2000-11-10
;  PRIOR APPLICATION NUMBER: US 60/249,596
;  PRIOR FILING DATE: 2000-11-17
;  NUMBER OF SEQ ID NOS: 68
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Peptide
US-10-010-160-68

  Query Match             7.9%;  Score 113;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 0.00036;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    248 PNSSSVDKLAAALEHHHHHH 267
          ||||||||||||||||||||
Db      2 PNSSSVDKLAAALEHHHHHH 21


RESULT 3
US-09-724-676-83340
; Sequence 83340, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83340
;   LENGTH: 1510
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83340

  Query Match             7.3%;  Score 104.5;  DB 5;  Length 1510;
  Best Local Similarity   18.9%;  Pred. No. 0.68;
  Matches   62;  Conservative   47;  Mismatches  106;  Indels  113;  Gaps   12;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||          |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 166
           |       |  :|: :: |                        ::| |  |: : :| :
Db   1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy    167 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR-------- 209
             :|||         | | | |:| :    |||:|    :|   |  | |
Db   1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARA 1423

Qy    210 ----------------VDLTDKNIYSRDGM---------------------LILALTRK 231
                          :|:  :|| :  |                        |:: | |
Db   1424 FVRKTSIFIMDEATASIDMATENILQKVVMTAFADRTVVTIAHRVHTILSADLVIVLKR- 1482

Qy    232 GQESFNGQVPRDDEPAPNSSSVDKLAAA 259
                | :   |:|    |  | : |:
Db   1483 ------GAILEFDKPEKLLSRKDSVFAS 1504


RESULT 4
US-09-724-676A-83340
; Sequence 83340, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83340
;   LENGTH: 1510
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83340

  Query Match             7.3%;  Score 104.5;  DB 5;  Length 1510;
  Best Local Similarity   18.9%;  Pred. No. 0.68;
  Matches   62;  Conservative   47;  Mismatches  106;  Indels  113;  Gaps   12;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||          |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 166
           |       |  :|: :: |                        ::| |  |: : :| :
Db   1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy    167 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR-------- 209
             :|||         | | | |:| :    |||:|    :|   |  | |
Db   1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARA 1423

Qy    210 ----------------VDLTDKNIYSRDGM---------------------LILALTRK 231
                          :|:  :|| :  |                        |:: | |
Db   1424 FVRKTSIFIMDEATASIDMATENILQKVVMTAFADRTVVTIAHRVHTILSADLVIVLKR- 1482

Qy    232 GQESFNGQVPRDDEPAPNSSSVDKLAAA 259
                | :   |:|    |  | : |:
Db   1483 ------GAILEFDKPEKLLSRKDSVFAS 1504


RESULT 5
US-09-724-676-83341
; Sequence 83341, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83341
;   LENGTH: 1554
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83341

  Query Match             7.2%;  Score 103.5;  DB 5;  Length 1554;
  Best Local Similarity   20.8%;  Pred. No. 0.88;
  Matches   46;  Conservative   36;  Mismatches   80;  Indels   59;  Gaps    8;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
```

```
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 166
          |       |  :|: :: |                        ::| |  |: : :| :
Db   1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy    167 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWG 199
               :|||        | | | |:| :    |||:|   |
Db   1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGQLEG 1404


RESULT 6
US-09-724-676A-83341
; Sequence 83341, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83341
;   LENGTH: 1554
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83341

  Query Match             7.2%;  Score 103.5;  DB 5;  Length 1554;
  Best Local Similarity   20.8%;  Pred. No. 0.88;
  Matches   46;  Conservative   36;  Mismatches   80;  Indels   59;  Gaps    8;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGTQGLRFNLWSSESAAWVG----------------------QFDESKLPLFQFINWV 166
          |       |  :|: :: |                        ::| |  |: : :| :
Db   1315 HG-------LLKTEAESYEGLLAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL 1367

Qy    167 KVYKYTPGQ--------GEGGSDFTLDWTDNFDTFDGSRWG 199
               :|||        | | | |:| :    |||:|   |
Db   1368 ----ISPGQKIGICGRTGSGKSSFSLAFFRMVDTFEGQLEG 1404


RESULT 7
US-09-455-294A-10
; Sequence 10, Application US/09455294A
; GENERAL INFORMATION:
;  APPLICANT: Bannon, Gary A.
;  APPLICANT:  Burks, Wesley A.
;  APPLICANT:  Caplan, Michael J.
;  APPLICANT:  Sampson, Hugh
;  APPLICANT:  Sosin, Howard
;  TITLE OF INVENTION: Peptide Antigens
;  FILE REFERENCE: 2002834-0004
;  CURRENT APPLICATION NUMBER: US/09/455,294A
;  CURRENT FILING DATE:  1999-12-06
;  PRIOR APPLICATION NUMBER: PCT/US96/15222
;  PRIOR FILING DATE: 1996-09-23
;  PRIOR APPLICATION NUMBER: 08/717,933
;  PRIOR FILING DATE: 1996-09-23
;  PRIOR APPLICATION NUMBER: 09/141,220
;  PRIOR FILING DATE: 1998-08-27
;  PRIOR APPLICATION NUMBER: 60/074,590
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: 60/074,624
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: 60/074,633
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: 60/073,283
;  PRIOR FILING DATE: 1998-01-31
;  NUMBER OF SEQ ID NOS: 110
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 526
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Modified Ara h
;   OTHER INFORMATION: 3 protein from Arachis hypogaea, IgE-binding sites
;   OTHER INFORMATION: disrupted
US-09-455-294A-10

  Query Match             7.0%;  Score 100.5;  DB 5;  Length 526;
  Best Local Similarity   54.1%;  Pred. No. 0.38;
  Matches   20;  Conservative    4;  Mismatches   12;  Indels    1;  Gaps    1;

Qy    231 KGQESFNGQVPRDDEPAPNSSSVDKLAAALEHHHHHH 267
          |    |  .||   : :| : :||||||||||||||
Db    491 KNNNPFKFFVP-PSQQSPRAVAVDKLAAALEHHHHHH 526


RESULT 8
US-09-724-676-83343
; Sequence 83343, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83343
;   LENGTH: 1558
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83343

  Query Match             6.9%;  Score 99.5;  DB 5;  Length 1558;
  Best Local Similarity   18.2%;  Pred. No. 2;
  Matches   67;  Conservative   44;  Mismatches  111;  Indels  147;  Gaps   13;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGT--------QGL-------RFNLWSSESAAWVG---------------------- 150
          |          :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                   ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    178 GGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR----------------------VDL 212
          | | |:| :    |||:|    :|   |  | |                      :|:
Db   1431 GKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARAFVRKTSIFIMDEATASIDM 1490

Qy    213 TDKNIYSRDGM----------------------LILALTRKGQESFNGQVPRDDEPAPNS 250
            :||  :  |                      |:: | |       | :   |:|
Db   1491 ATENILQKVVMTAFADRTVVTIAHRVHTILSADLVIVLKR-------GAILEFDKPEKLL 1543
```

```
Qy    251 SSVDKLAAA 259
          |  | : |:
Db   1544 SRKDSVFAS 1552


RESULT 9
US-09-724-676A-83343
; Sequence 83343, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83343
;   LENGTH: 1558
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83343

  Query Match             6.9%;  Score 99.5;  DB 5;  Length 1558;
  Best Local Similarity   18.2%;  Pred. No. 2;
  Matches   67;  Conservative   44;  Mismatches  111;  Indels  147;  Gaps   13;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  | :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 150
           |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                   ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    178 GGSDFTLDWTDNFDTFDGSRWGKGDWTF-DGNR------------------------VDL 212
          | | |:| :    |||:|    :|   |  | |                        :|:
Db   1431 GKSSFSLAFFRMVDTFEGHIITEGGENFSQGQRQLFCLARAFVRKTSIFIMDEATASIDM 1490

Qy    213 TDKNIYSRDGM----------------------LILALTRKGQESFNGQVPRDDEPAPNS 250
            :||  :  |                      |:: | |       |  :   |:|
Db   1491 ATENILQKVVMTAFADRTVVTIAHRVHTILSADLVIVLKR-------GAILEFDKPEKLL 1543

Qy    251 SSVDKLAAA 259
          |  | : |:
Db   1544 SRKDSVFAS 1552


RESULT 10
US-09-724-676-83344
; Sequence 83344, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83344
;   LENGTH: 1602
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83344

  Query Match             6.8%;  Score 98.5;  DB 5;  Length 1602;
  Best Local Similarity   19.5%;  Pred. No. 2.6;
  Matches   51;  Conservative   33;  Mismatches   85;  Indels   93;  Gaps    9;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  | :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 150
           |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                   ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    178 GGSDFTLDWTDNFDTFDGSRWG 199
          | | |:| :    |||:|    |
Db   1431 GKSSFSLAFFRMVDTFEGQLEG 1452


RESULT 11
US-09-724-676A-83344
; Sequence 83344, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83344
;   LENGTH: 1602
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83344

  Query Match             6.8%;  Score 98.5;  DB 5;  Length 1602;
  Best Local Similarity   19.5%;  Pred. No. 2.6;
  Matches   51;  Conservative   33;  Mismatches   85;  Indels   93;  Gaps    9;

Qy     12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
          | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db   1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy     71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
          : ::   |    ::  | :|        || |    ||: | |      : : | |  :
Db   1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy    131 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 150
           |         :||       |    | |   |||
Db   1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy    151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                   ::| |  |: : :| :     :|||        |
Db   1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy    178 GGSDFTLDWTDNFDTFDGSRWG 199
          | | |:| :    |||:|    |
Db   1431 GKSSFSLAFFRMVDTFEGQLEG 1452


RESULT 12
US-09-724-676-83342
; Sequence 83342, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
```

```
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83342
;   LENGTH: 1629
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676-83342

  Query Match             6.8%;  Score 97.5;  DB 5;  Length 1629;
  Best Local Similarity   19.4%;  Pred. No. 3.3;
  Matches   50;  Conservative   33;  Mismatches   82;  Indels   93;  Gaps    9;

Qy      12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     131 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 150
            |         :||       |    | |   |||
Db    1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy     151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                    ::| |  |: : :| :     :|||        |
Db    1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy     178 GGSDFTLDWTDNFDTFDG 195
           | | |:| :    |||:|
Db    1431 GKSSFSLAFFRMVDTFEG 1448


RESULT 13
US-09-724-676A-83342
; Sequence 83342, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83342
;   LENGTH: 1629
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-724-676A-83342

  Query Match             6.8%;  Score 97.5;  DB 5;  Length 1629;
  Best Local Similarity   19.4%;  Pred. No. 3.3;
  Matches   50;  Conservative   33;  Mismatches   82;  Indels   93;  Gaps    9;

Qy      12 LYTLEEVQY-GKFEARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70
           | |:   :|  :|: ::     |  ::|:||           |:|| :| :|
Db    1210 LTTIRAFRYEARFQQKLLEYTDSNNIASLFL-------TAANRWLEVRMEYIG------- 1255

Qy      71 SNIITGKAGAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNL 130
           : ::   |    ::  |  :|        || |    ||: | |      : : | |  :
Db    1256 ACVVLIAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNYLNWMVRNLADMELQLGAVKRI 1314

Qy     131 TGT--------QGL-------RFNLWSSESAAWVG------------------------ 150
            |         :||       |    | |   |||
Db    1315 HGLLKTEAESYEGLLGERLRERGGEESKEECVWVGGHKGAWGWGGTFGYSCGPCLVLSPA 1374

Qy     151 ------------------------QFDESKLPLFQFINWVKVYKYTPGQ--------GE 177
                                    ::| |  |: : :| :     :|||        |
Db    1375 RPPAPSLIPKNWPDQGKIQIQNLSVRYDSSLKPVLKHVNAL----ISPGQKIGICGRTGS 1430

Qy     178 GGSDFTLDWTDNFDTFDG 195
           | | |:| :    |||:|
Db    1431 GKSSFSLAFFRMVDTFEG 1448


RESULT 14
US-10-141-531-59
; Sequence 59, Application US/10141531
; GENERAL INFORMATION:
;  APPLICANT: Briggs, Steven P.
;  APPLICANT:  Dalmia, Bipin K.
;  APPLICANT:  del Val, Greg
;  APPLICANT:  Desjarlais, John R.
;  APPLICANT:  Heifetz, Peter
;  APPLICANT:  Luginbuhl, Peter
;  APPLICANT:  Muchhal, Umesh
;  TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
;  FILE REFERENCE: A-71457-2/RFT/RMS/RMK
;  CURRENT APPLICATION NUMBER: US/10/141,531
;  CURRENT FILING DATE:  2002-05-06
;  PRIOR APPLICATION NUMBER: US 60/370,609
;  PRIOR FILING DATE: 2002-04-05
;  PRIOR APPLICATION NUMBER: US 60/289,029
;  PRIOR FILING DATE: 2001-05-04
;  NUMBER OF SEQ ID NOS: 208
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59
;   LENGTH: 481
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-141-531-59

  Query Match             6.7%;  Score 97;  DB 6;  Length 481;
  Best Local Similarity   20.7%;  Pred. No. 0.71;
  Matches   56;  Conservative   30;  Mismatches   88;  Indels   96;  Gaps   10;

Qy      79 GAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRKTEGGQVSNLTGTQGLRF 138
           || |  ::    :|  |  :::  :|   :  |  : | :|:    | ||:|    || |
Db     227 GASKIMQQRALSNPKIDVIWNSSVVEAYGDGERDVLGGLKVKNVVTGDVSDLK-VSGLFF 285

Qy     139 NLWSSESAAWVGQFDESKLPLFQFINWVKVYKYTPGQGEGGSDFTLDWTDNF-------- 190
            :    :  ::    :| :    |:       ||       |   : : | | |
Db     286 AIGHEPATKFLDGGVDSAMEEANFLT-----KY------GSKVYIIHWVDAFGASKIMQQ 334

Qy     191 --------DTFDGS----RWGKGDWTFDGNRVDLTDKNIYSRD------GMLILALTRKG 232
                   |    |      :| |:    |    |  ||: : |        |  |:  :
Db     335 RALSNPKIDVIWNSSVVEAYGDGERDVLGG---LKVKNVVTGDVSDLKVSGLFFAIGHEP 391

Qy     233 QESF-NGQVPRDDE------PAPNSSS-------------------------------- 252
             | :| |  | :       |    :|
Db     392 ATKFLDGGVELDSDGYVVTKPGTTQTSVPGVFAAGDVQDKKYRQAITAAGTGCMAALDAE 451

Qy     253 ---------------VDKLAAALEHHHHHH 267
                           |||||||||||||||
Db     452 HYLQEIGSQQGKSDGVDKLAAALEHHHHHH 481


RESULT 15
US-09-424-705B-4
; Sequence 4, Application US/09424705B
; GENERAL INFORMATION:
;  APPLICANT: LITTLE, MELVYN
;  APPLICANT:  KIPRIYANOV, SERGEY
;  APPLICANT:  MOLDENHAUER, GERHARD
;  APPLICANT:  DEUTSCHES KREBSFORSCHUNGSZEUTRUM
;  TITLE OF INVENTION: MUTATED OKT3 ANTIBODY
;  FILE REFERENCE: 035280047US00
;  CURRENT APPLICATION NUMBER: US/09/424,705B
;  CURRENT FILING DATE:  2000-06-02
;  PRIOR APPLICATION NUMBER: PCT/DE98/01409
;  PRIOR FILING DATE: 1998-05-22
```

```
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 290
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-424-705B-4

  Query Match             6.5%;  Score 94;  DB 5;  Length 290;
  Best Local Similarity   20.4%;  Pred. No. 0.67;
  Matches   66;  Conservative   43;  Mismatches   91;  Indels  124;  Gaps   18;

Qy      25 ARMKMAAASGTVSSMFLYQNGSEIADGRP--------------------WVE----VDI 59
           | : : ||   :: : | |:|:|:|  ||                    ||:      :
Db      10 AGLLLLAAQPAMAQVQLQQSGAELA--RPGASVKMSCKASGYTFTRYTMHWVKQRPGQGL 67

Qy      60 EVLGK-NPGSFQSNI---ITGKAGAQKTSEKHHAVS-------PAADQAFH--------T 100
           | :|  ||    :|       |  |  |::|  : :        : | | :
Db      68 EWIGYINPSRGYTNYNQKFKDKDKATLTTDKSSSTAYMQLSSLTSEDSAVYYCARYYDDH 127

Qy     101 YGLEWTPNYVRWTVDGQEVRKTEGGQV------SNLTGTQGLR--FNLWSSESAAWVGQF 152
           | |::       ||   :     || :      ::|  : | |   :  :|:|  : |
Db     128 YSLDYWGQGTTLTVSSAKTTPKLGGDILLTQTPASLAVSLGQRATISCKASQSVDYDGD- 186

Qy     153 DESKLPLFQFINWVKVYKYTPGQ--------------------GEG-GSDFTLDW--TD 188
                 ::||   |:  |||                     | | |:||||:    :
Db     187 --------SYLNW---YQQIPGQPPKLLIYDASNLVSGIPPRFSGSGSGTDFTLNIHPVE 235

Qy     189 NFD--TFDGSRWGKGDWTFDGNRVDLTDKNIYSRDGMLILALTRKGQESFNGQVPRDDEP 246
             |  |:   :  :  ||| |                              :: | |
Db     236 KVDAATYHCQQSTEDPWTFGGGT-------------------------KLEIKRADAA 268

Qy     247 APNSSSVDKLAAALE---HHHHHH 267
           |  : |  || :  :   ||||||
Db     269 A--AGSEQKLISEEDLNSHHHHHH 290


Search completed: January  9, 2003, 12:14:13
Job time : 13.7431 secs
```